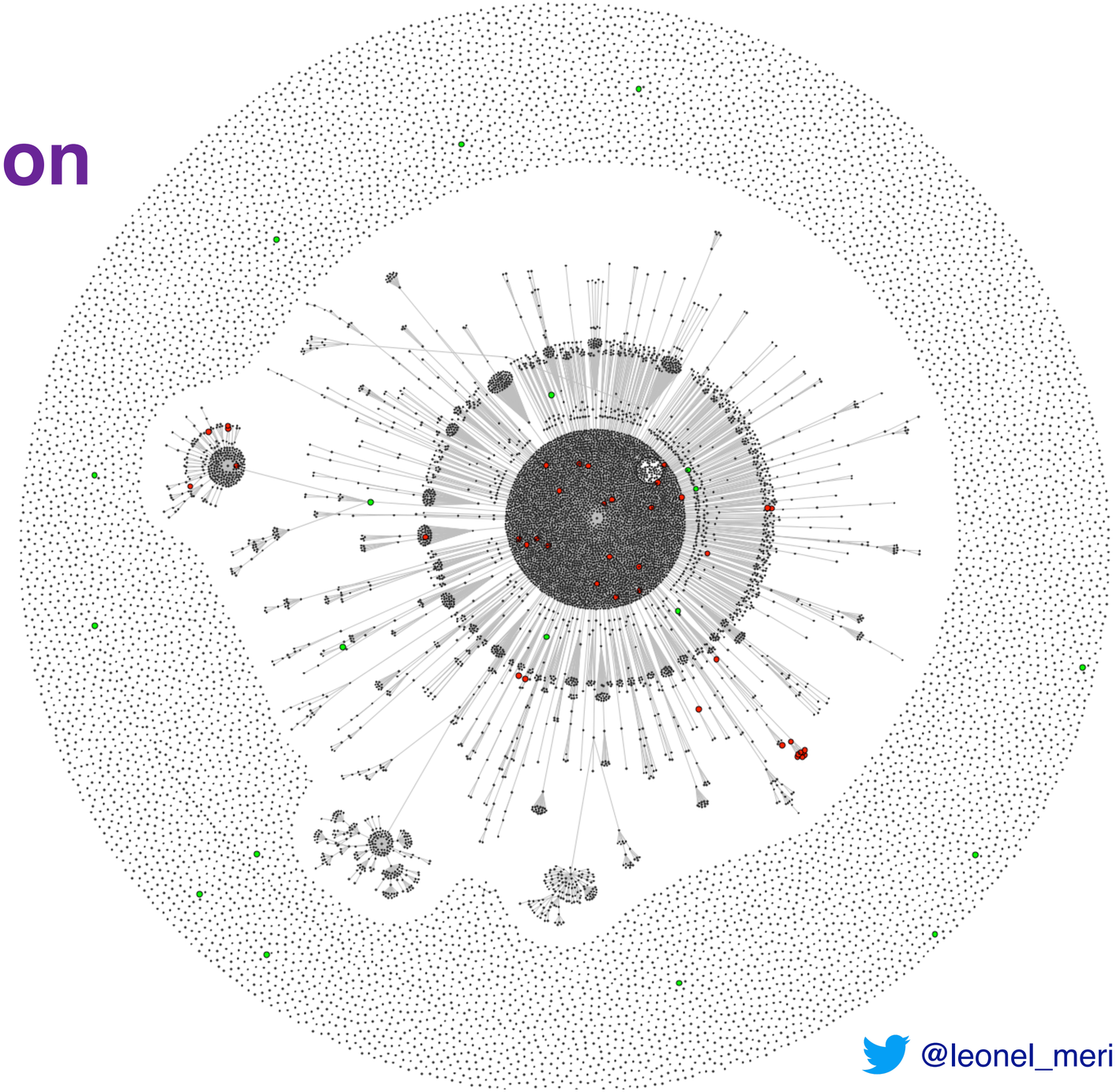
Software Visualization



Leonel Merino

Visualization Research Center (VISUS)
University of Stuttgart

leonel.merino@visus.uni-stuttgart.de

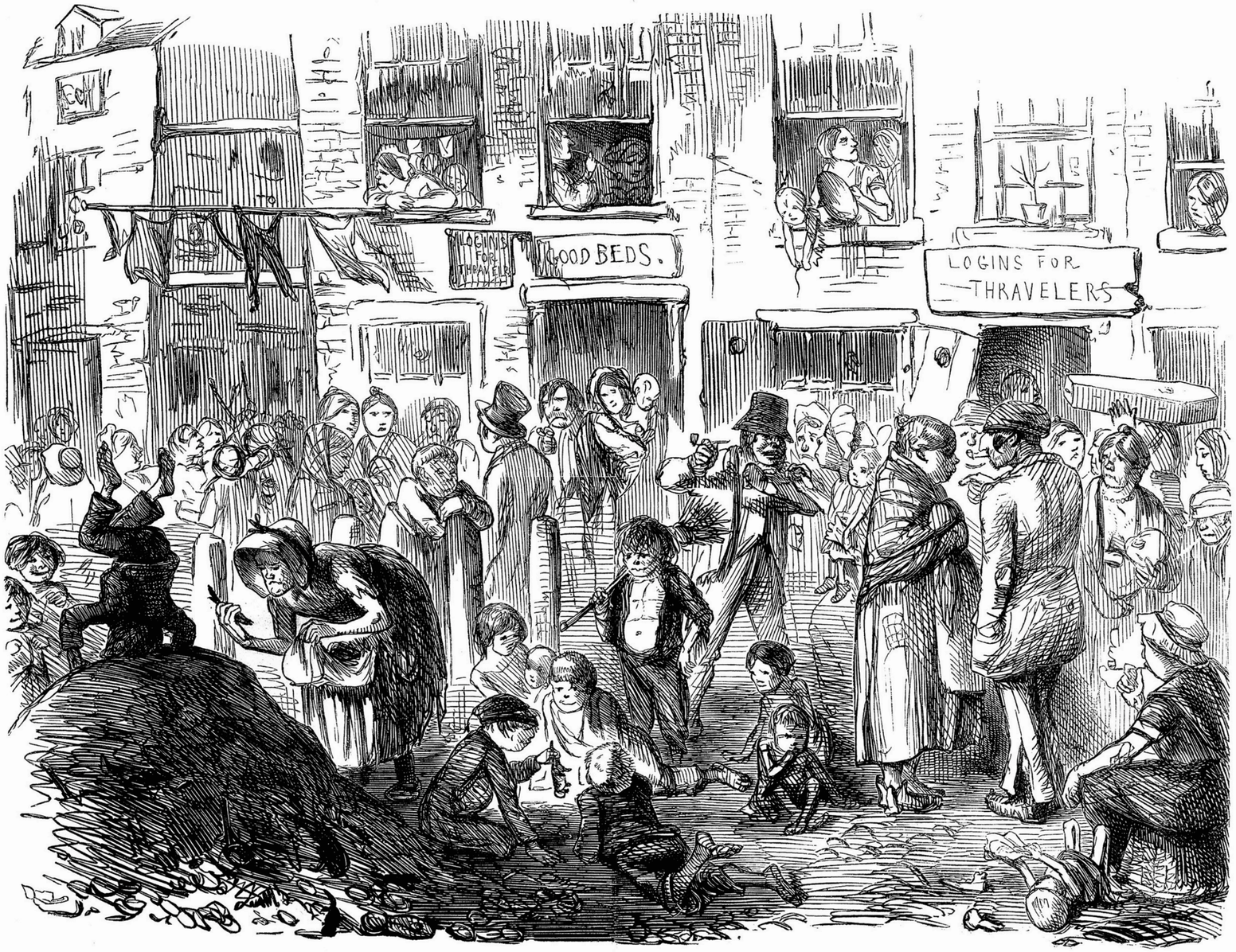
@leonel_merino

Roadmap

- Motivation
- Visual Perception
- Information Visualization
- Software Visualization

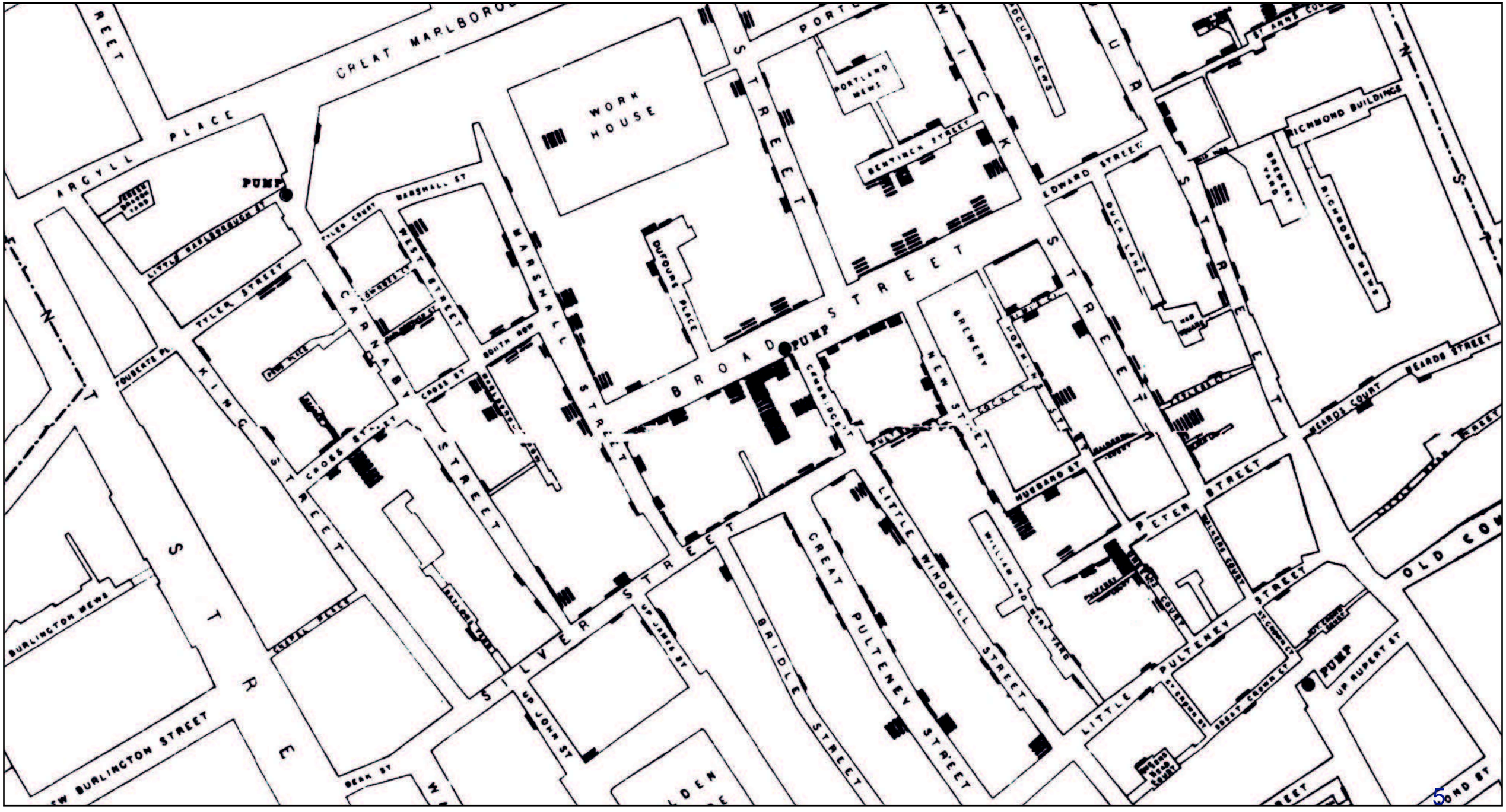
Roadmap

- **Motivation**
- Visual Perception
- Information Visualization
- Software Visualization

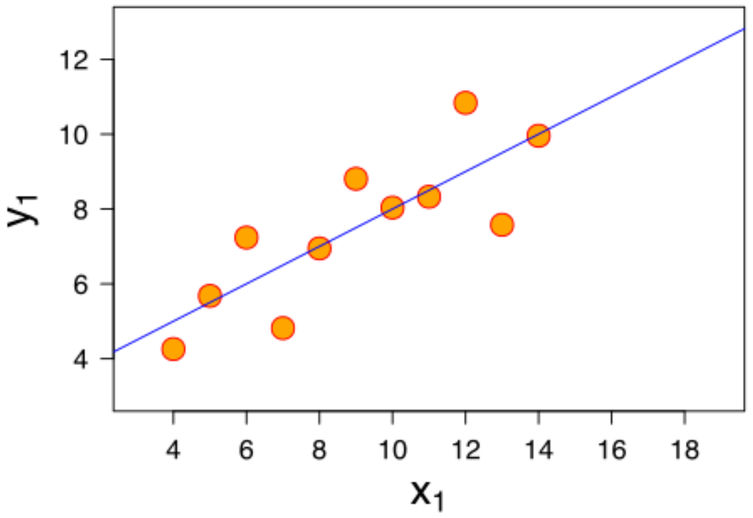


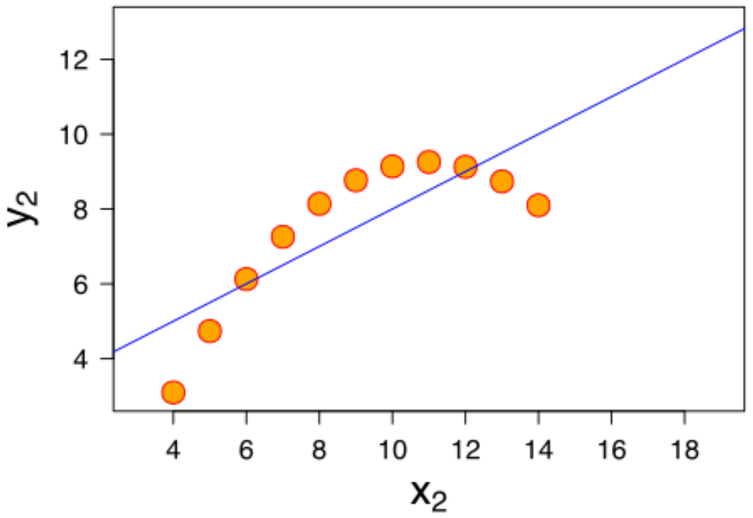
A COURT FOR KING CHOLERA.

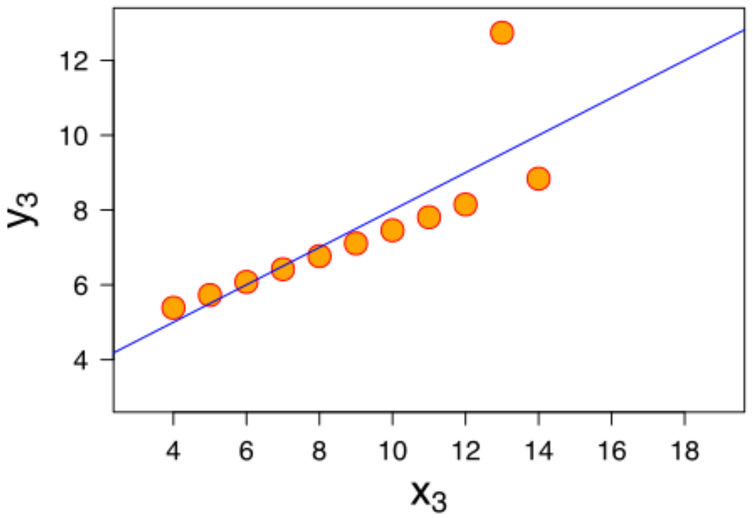
SEPTEMBER 25, 1852.

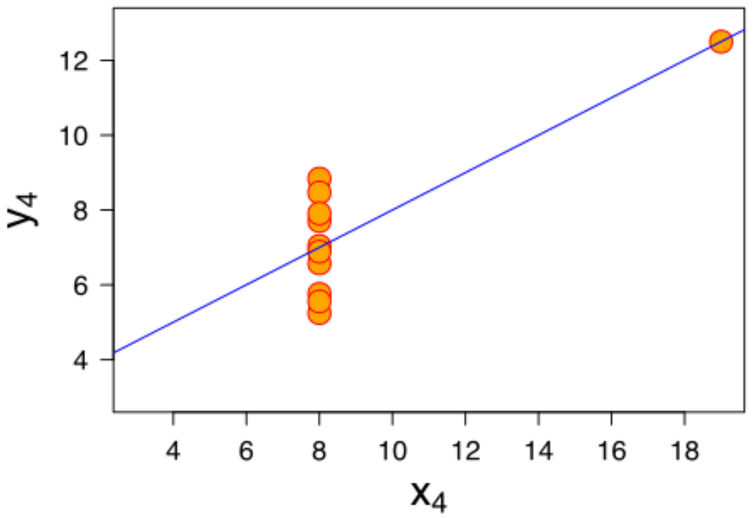

GREAT MARLBOROUGH STREET
ARGYLL PLACE
PUMP
LITTLE MARLBOROUGH ST
TYLER STREET
TYLER COURT
MARSHALL ST
WEST STREET
CROWNDALE
WORK HOUSE
DUFOURS PLACE
PORTLAND MEWS
BERWICK STREET
NICK
EDWARD STREET
DUCK LANE
RICHMOND BUILDINGS
BROAD STREET
PUMP
GREWERY
NEW STREET
HUSBAND ST
PETER STREET
KING STREET
CROSS STREET
FOUBERTS PL
BURLINGTON MEWS
BURLINGTON STREET
GREAT PULTENEY STREET
LITTLE WINDMILL STREET
BRIDLE STREET
LITTLE PULTENEY
PUMP
UP RUPERT ST
OLD COMPTON STREET
BEARDS COURT
BEARDS STREET
UP JAMES ST
UP JOHN ST
BEAK ST

Anscombe's Quartet





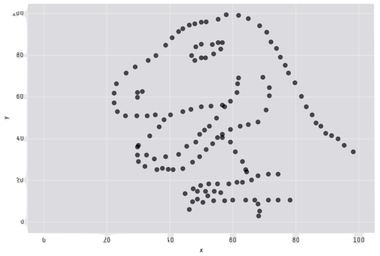




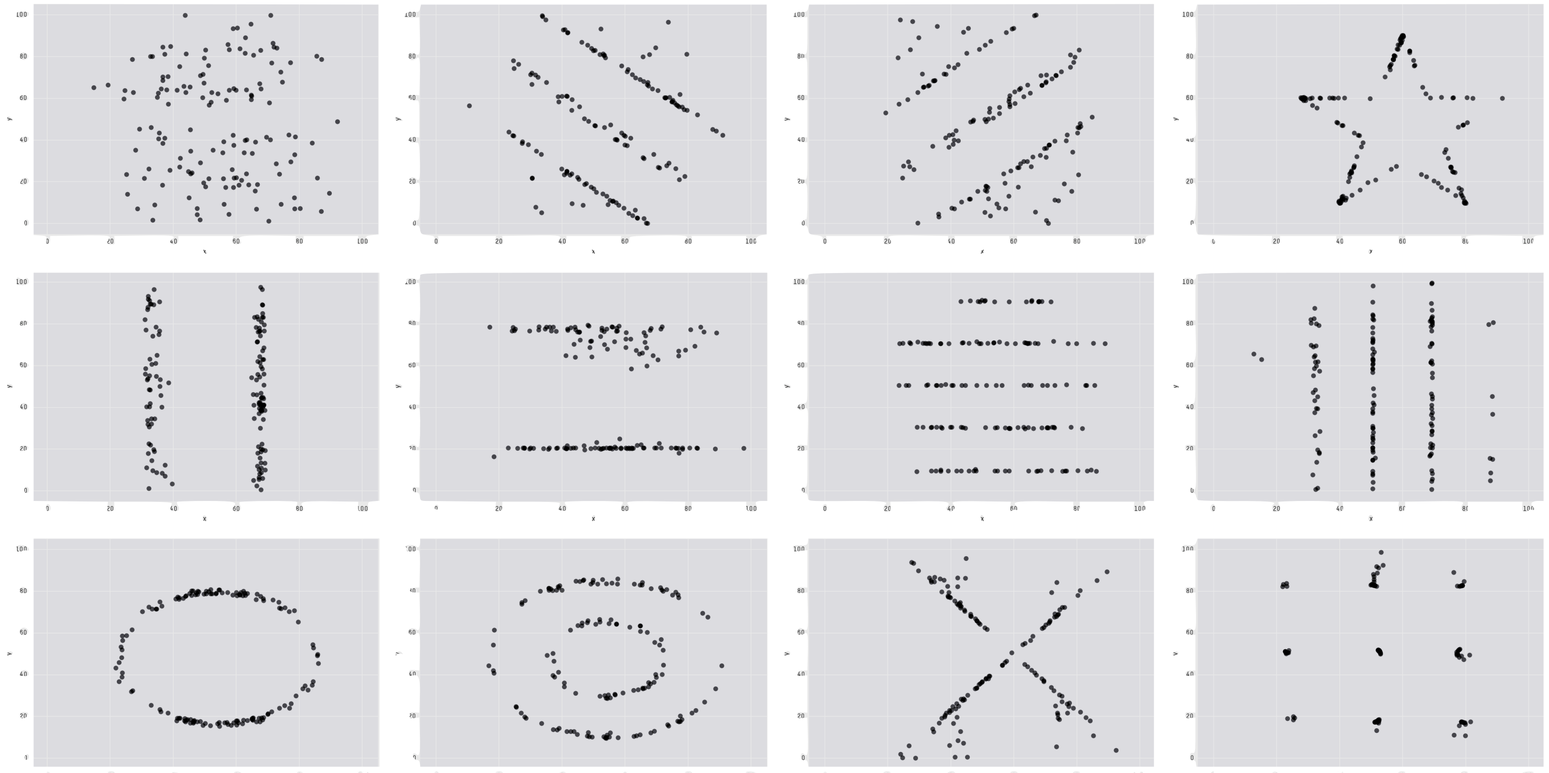
Number of observations $(n) = 11$
Mean of the x's $(\bar{x}) = 9.0$
Mean of the y's $(\bar{y}) = 7.5$
Regression coefficient (b_1) of y on $x = 0.5$
Equation of regression line: $y = 3 + 0.5\,x$
Sum of squares of $x - \bar{x} = 110.0$
Regression sum of squares $= 27.50$ (1 d.f.)
Residual sum of squares of $y = 13.75$ (9 d.f.)
Estimated standard error of $b_1 = 0.118$

F.J. Anscombe. Graphs in statistical analysis. *American Statistician*, 27(1):17–21, Feb. 1973

Datasaurus data set



X Mean: 54.26
Y Mean: 47.83
X SD : 16.76
Y SD : 26.93
Corr. : -0.06



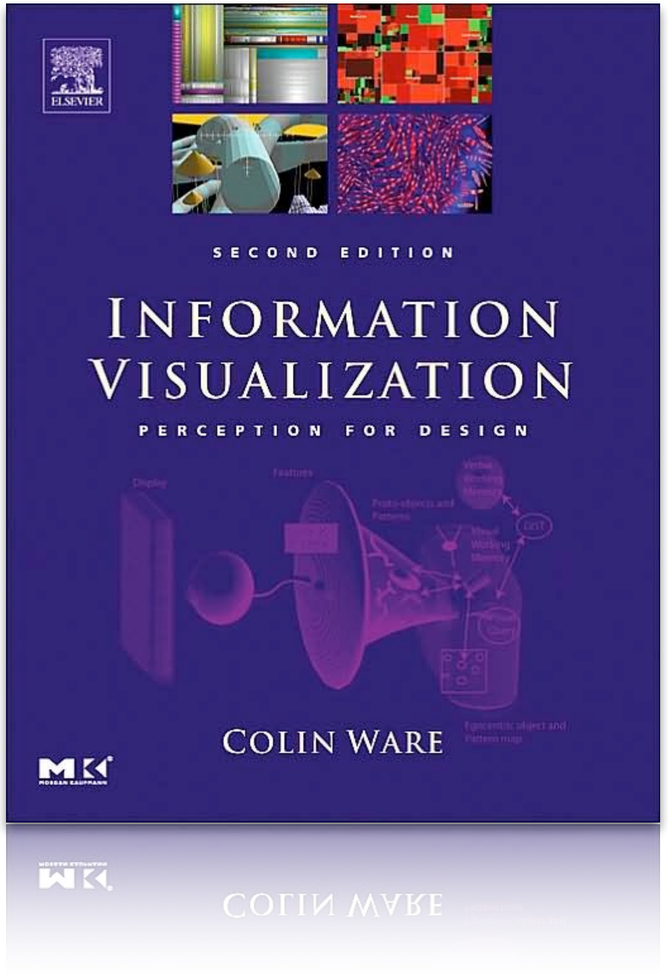
Matejka, J., & Fitzmaurice, G. (2017, May). Same stats, different graphs: generating datasets with varied appearance and identical statistics through simulated annealing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1290-1294). ACM.

Roadmap

- Motivation
- **Visual Perception**
- Information Visualization
- Software Visualization

Vision

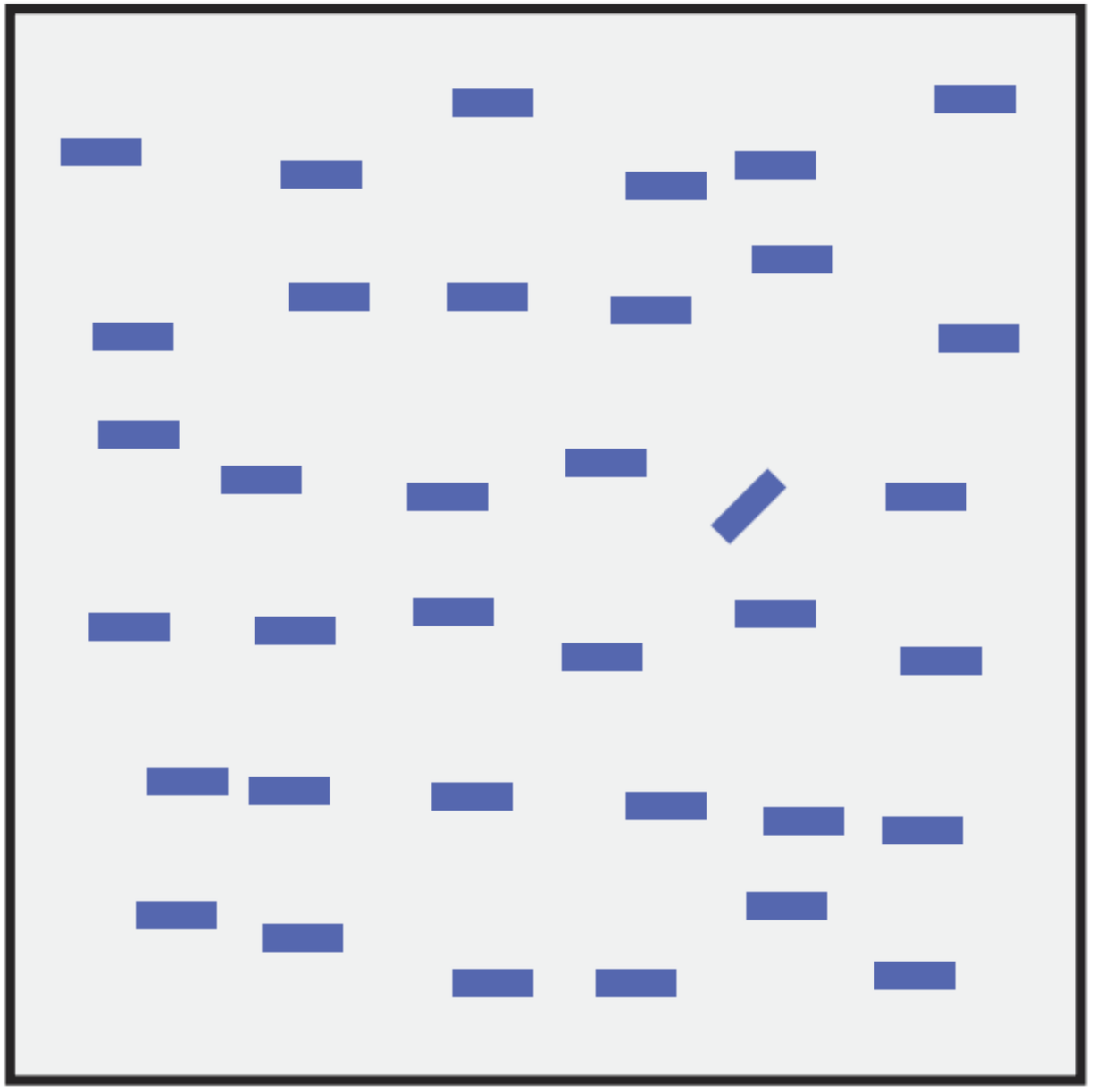
We acquire **more information through vision** than all the other senses combined



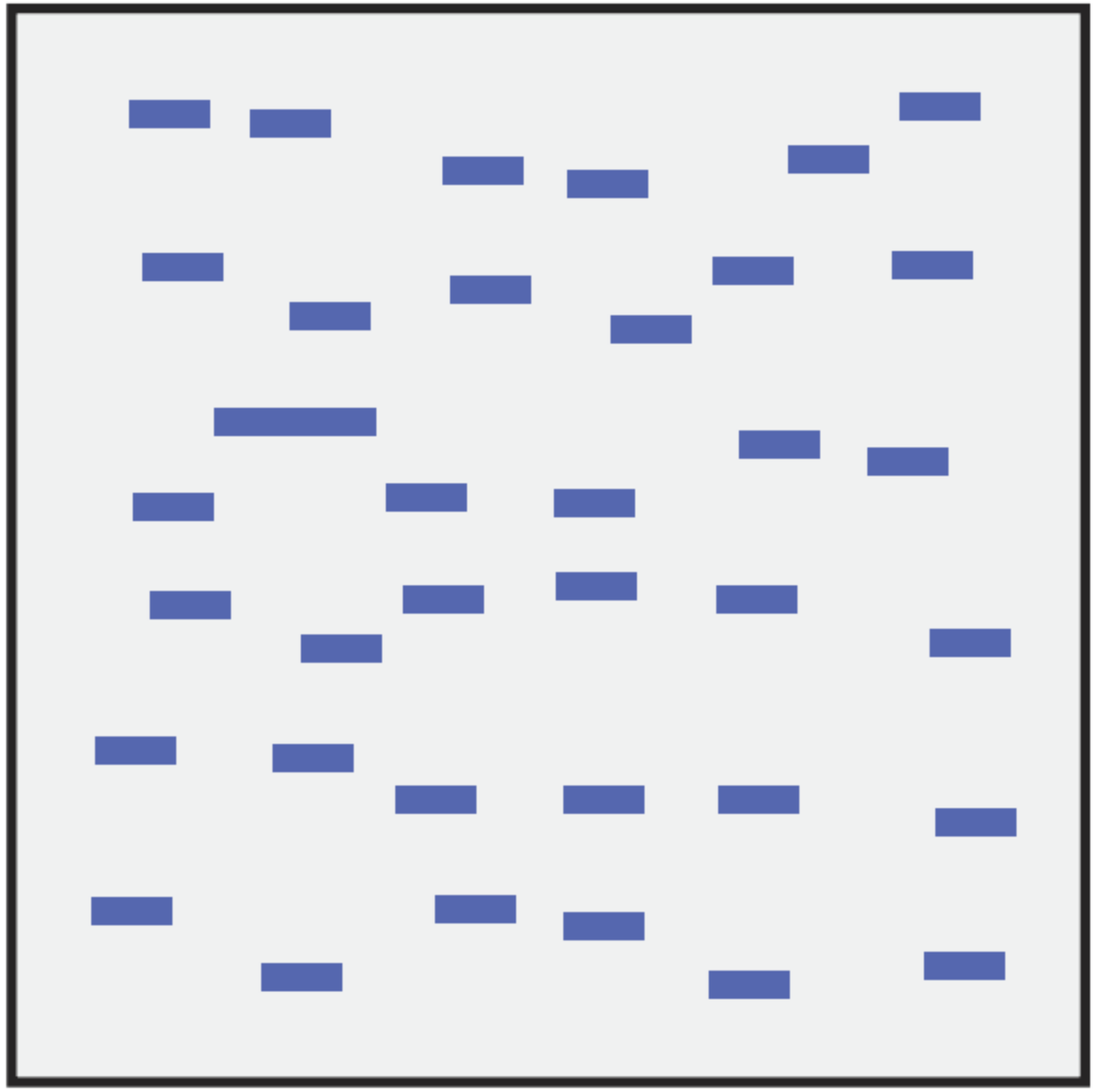
Pre-attentive Processing

*“Tasks that can be **performed** on large multi-element displays in less than **200-250 milliseconds** (msec) are considered **pre-attentive**.”*

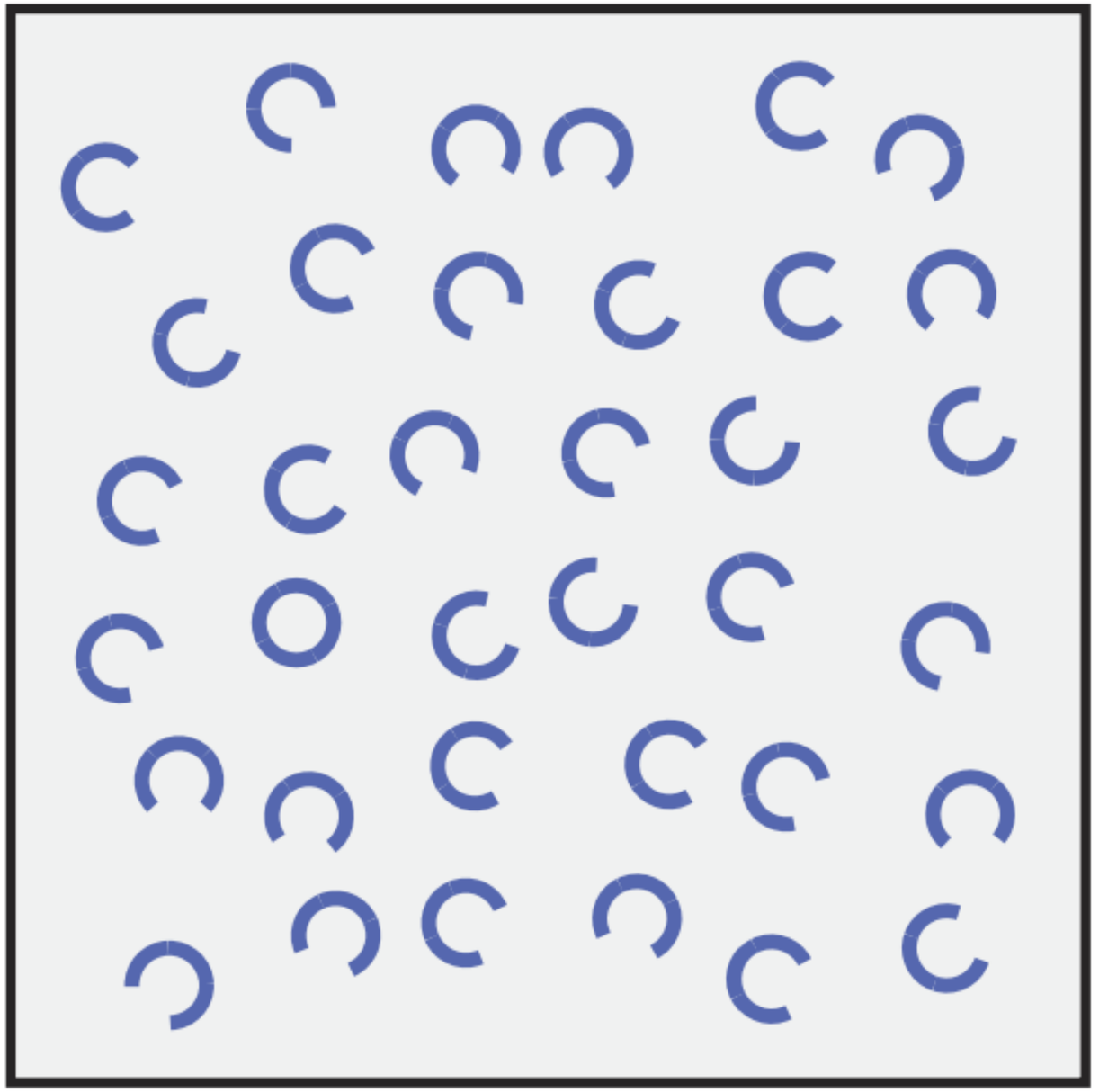
[Healy and Enns ’12]



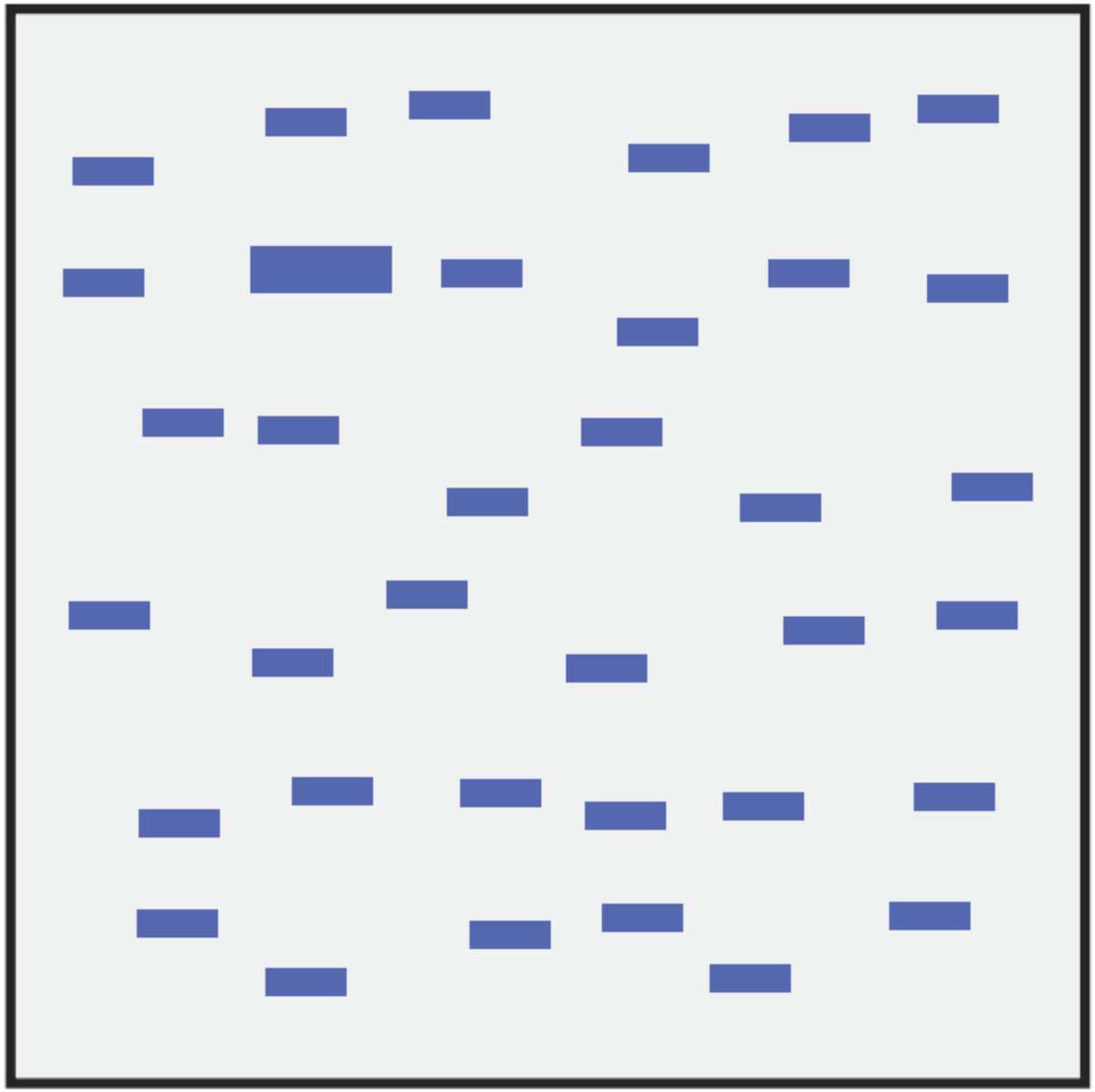
orientation



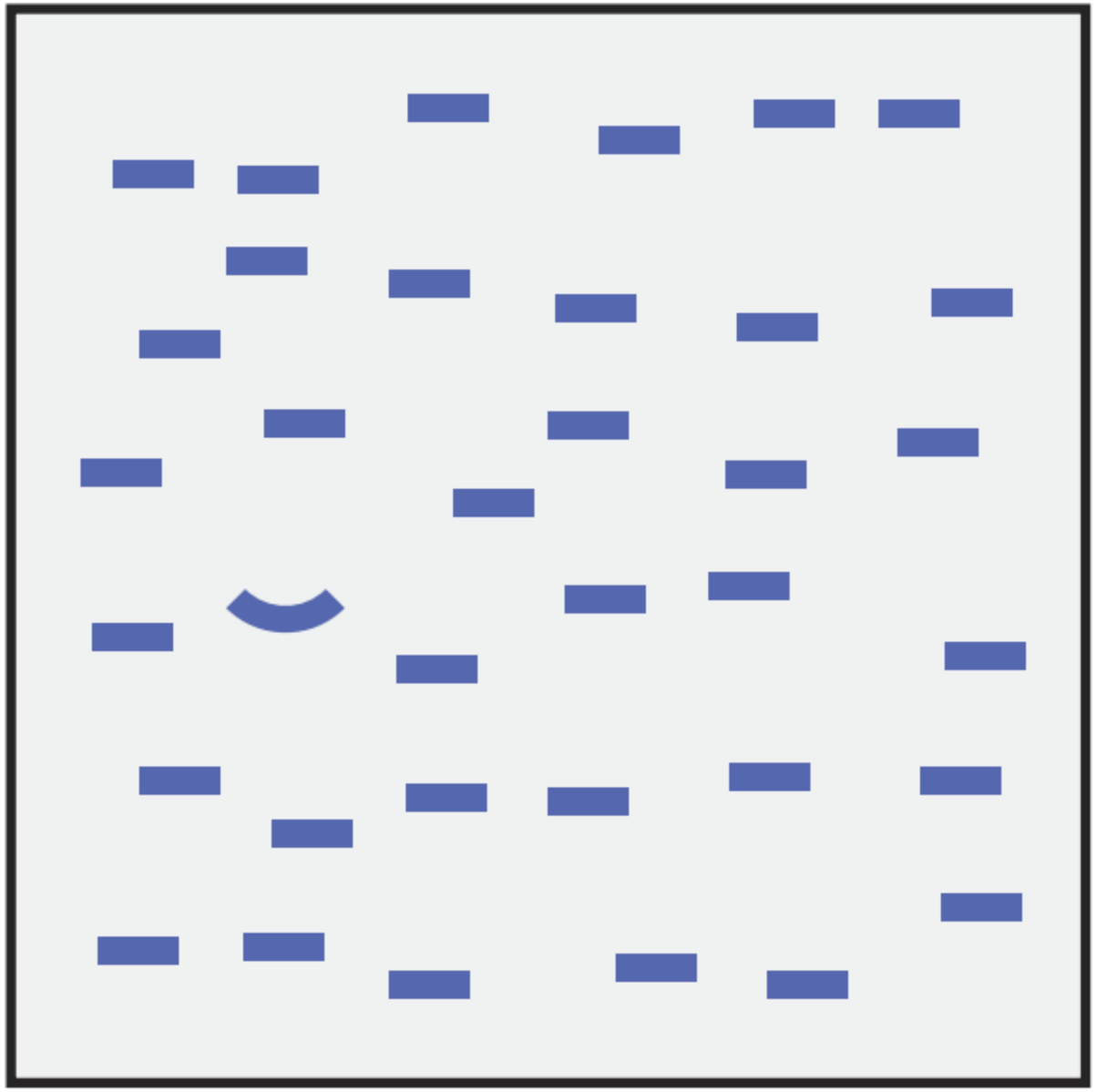
length



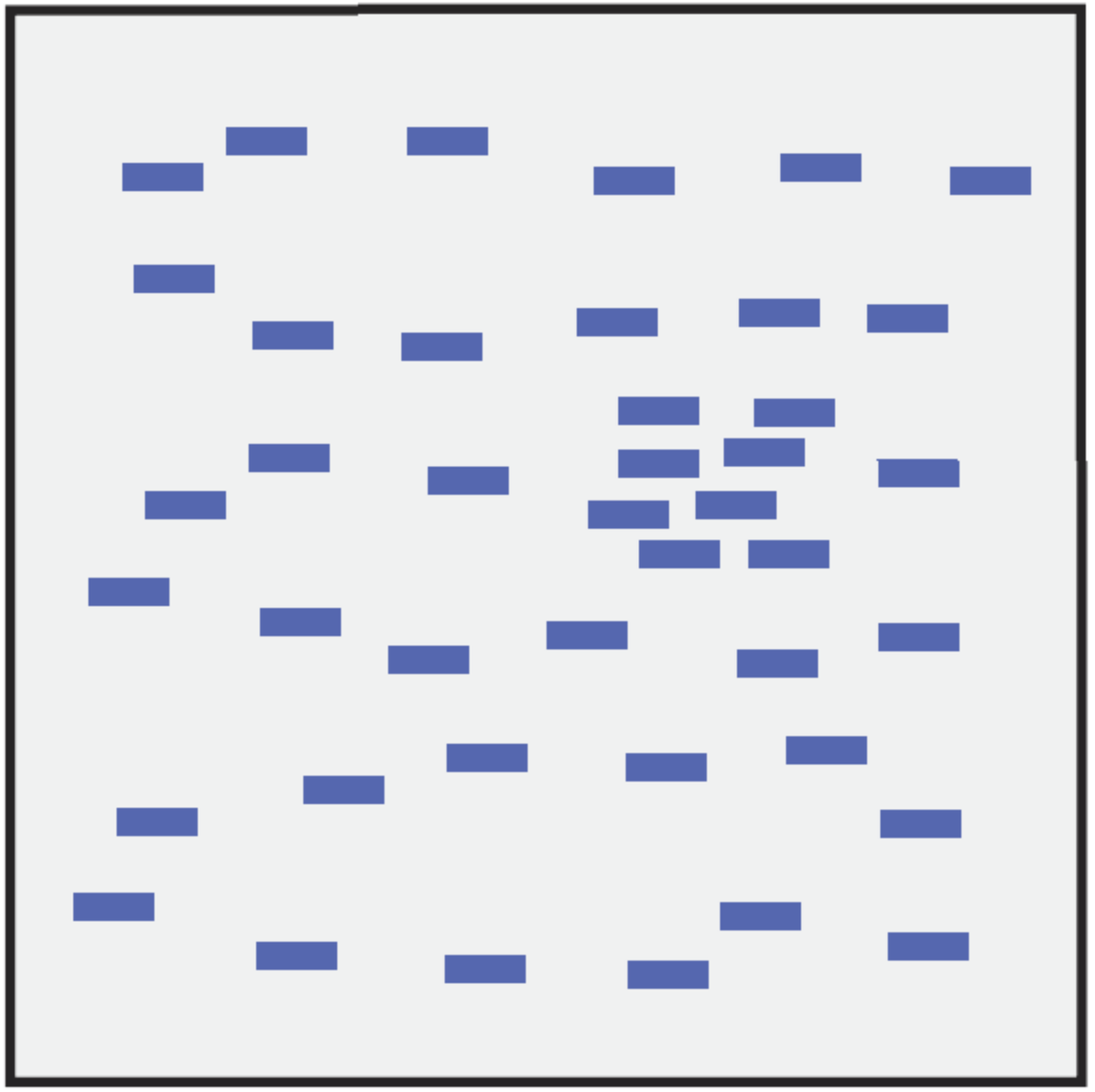
closure



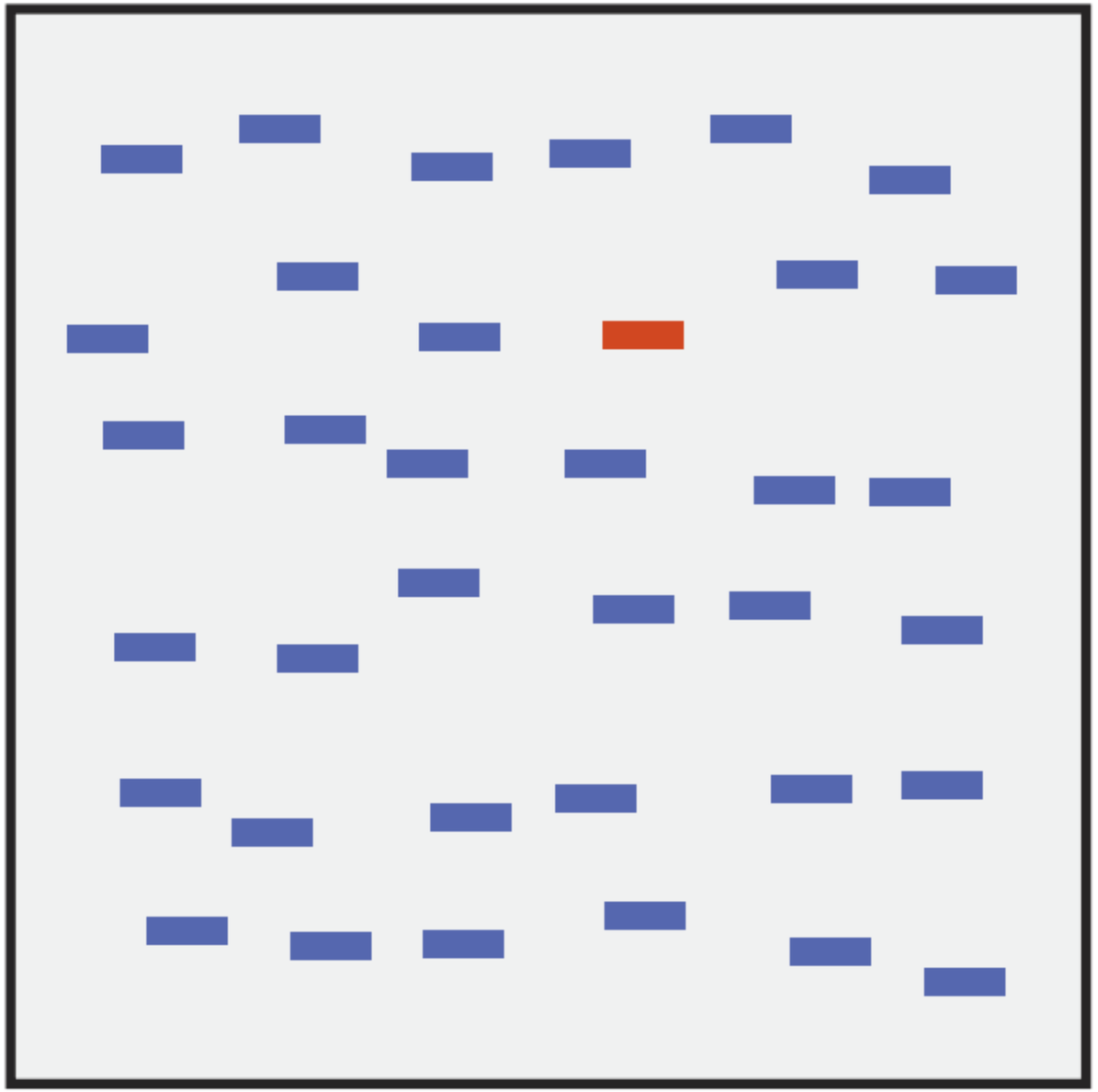
size



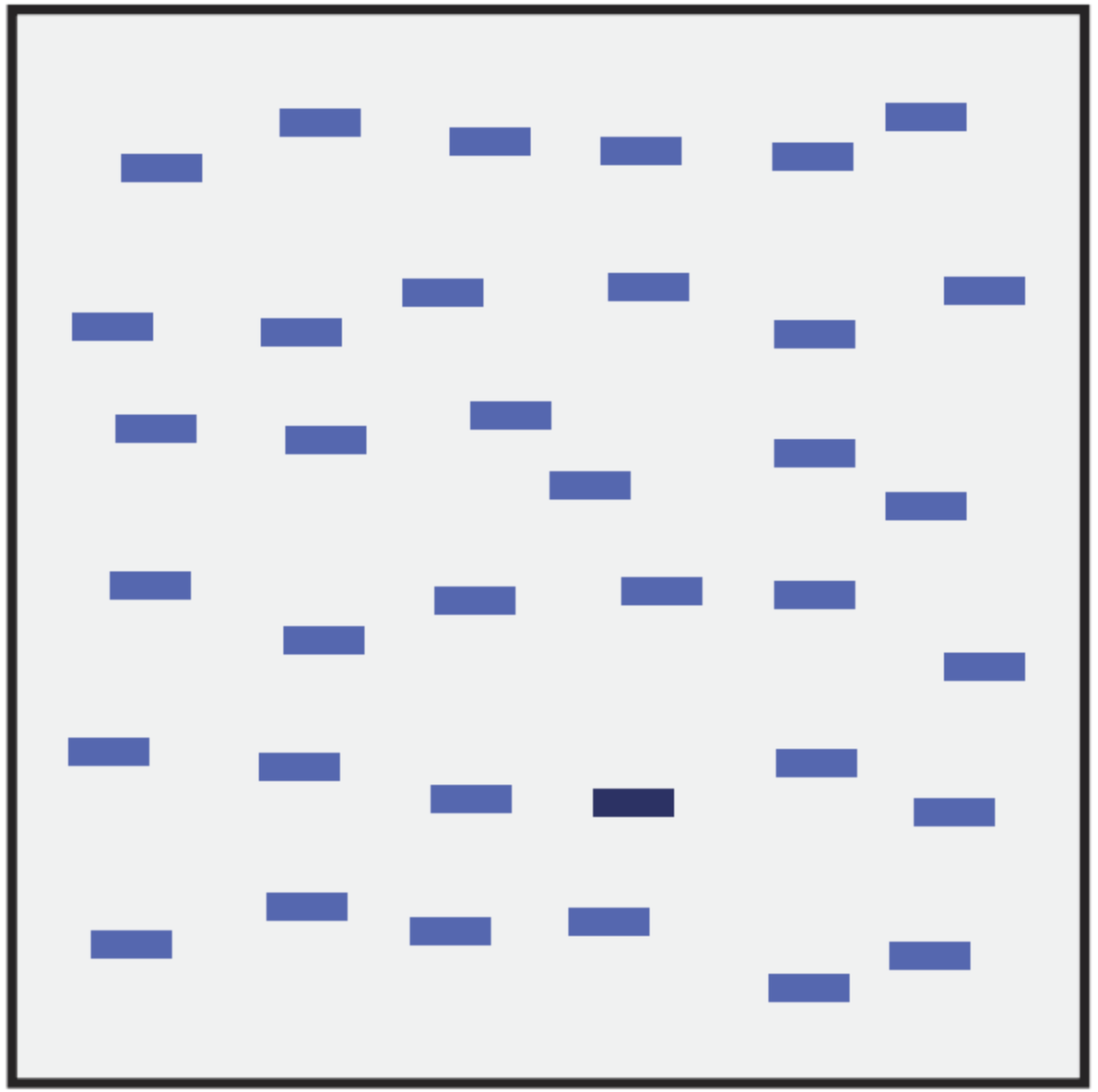
curvature



density

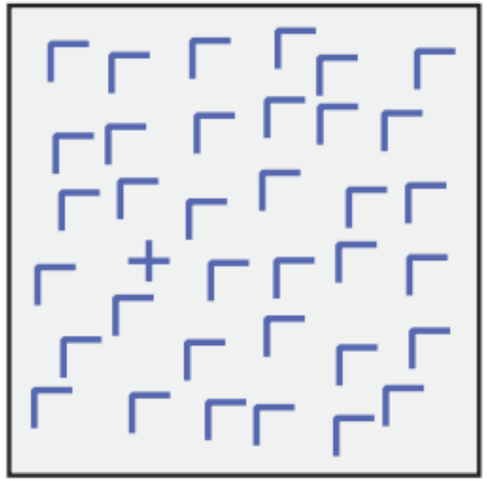


hue



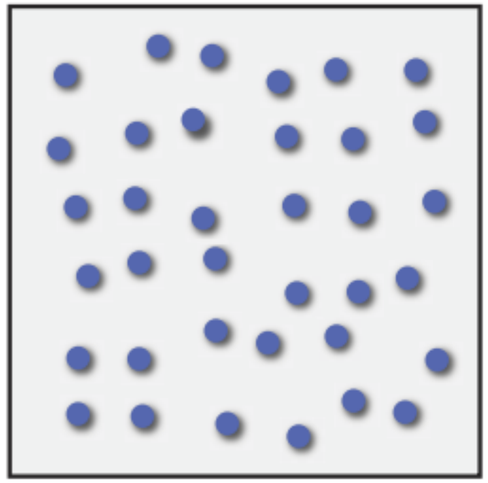
luminance

Preattentive Processing

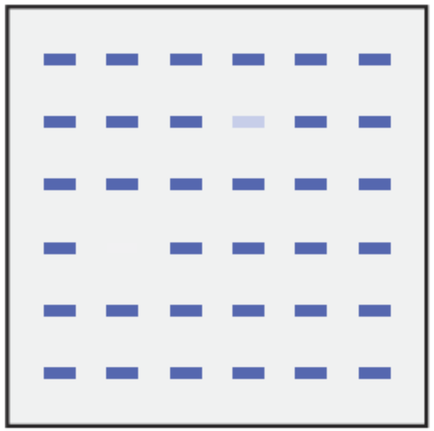


intersections

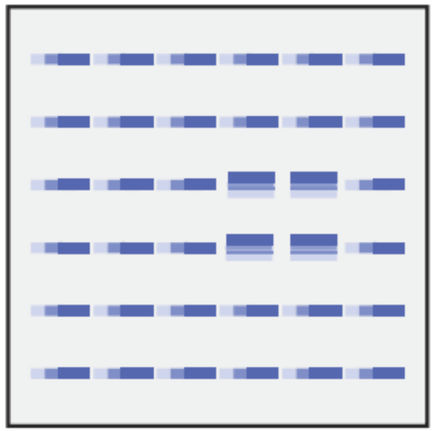
terminators



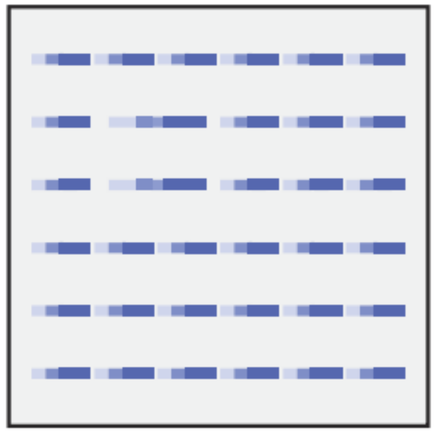
3D depth



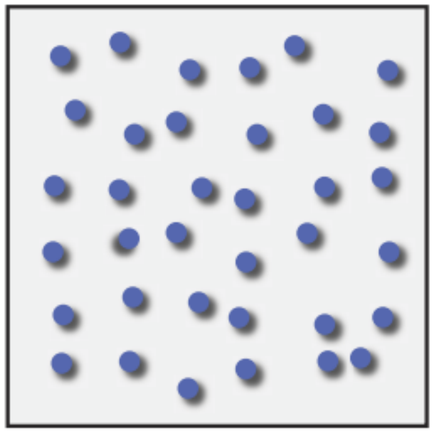
flicker



direction of motion

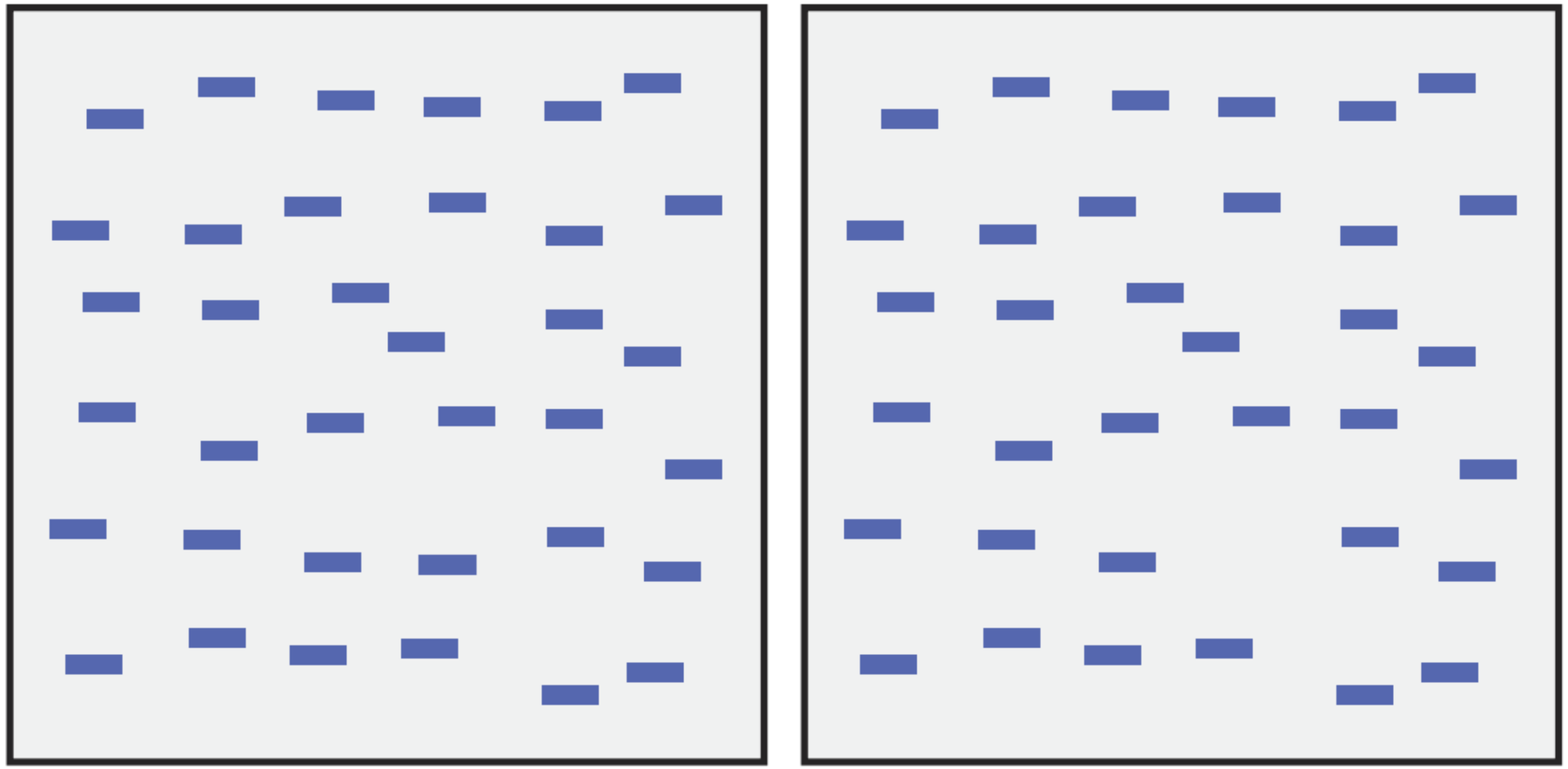


velocity of motion



lighting direction

Preattentive Processing



deadeye

Krekhov, Andrey, and Jens Krüger. "Deadeye: A Novel Preattentive Visualization Technique Based on Dichoptic Presentation." *IEEE transactions on visualization and computer graphics* (2018).

Pre-attentive Processing: Color

8789364082376403128764532984732984732094873290845389274-0329874-32874-2319847509834098340983240983204982323-0984903281453209481-083939394789658743659821

Pre-attentive Processing: Color

8789364082376403128764532984732984732094873290845 38
9274-0329874-32874-2319847509834098340983240983204 98
23-0984903281453209481-0839393947896587436598

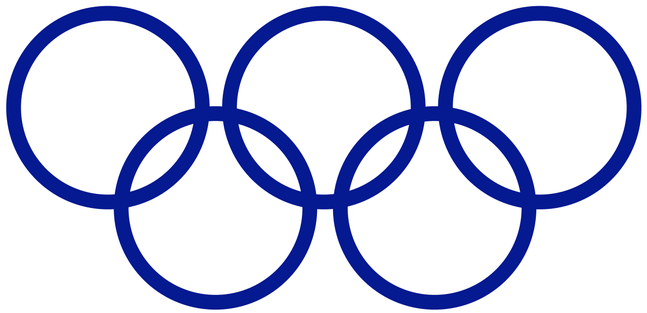
Gestalt Psychology

> The Gestalt Laws

1. Simplicity
2. Closure
3. Similarity
4. Proximity
5. Continuity

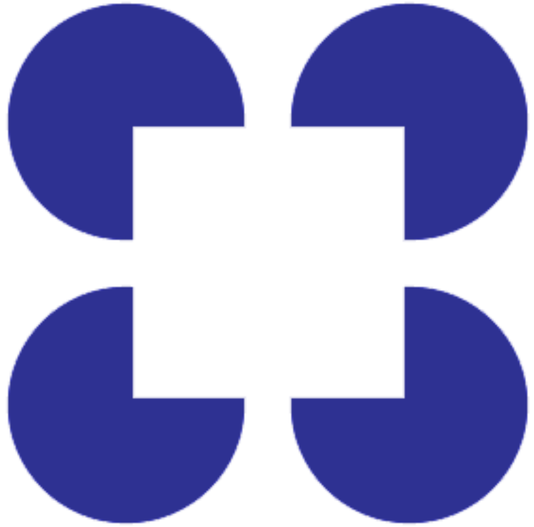
1. Law of Simplicity

> Reality is organized and reduced to the simplest form possible



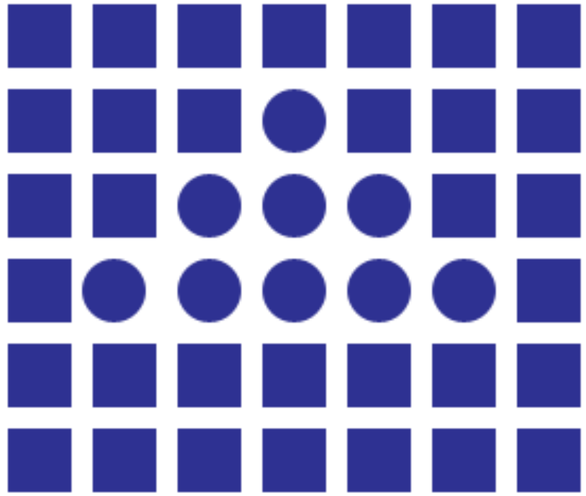
1. Law of Closure

> The mind completes missing parts so it can see a simple image

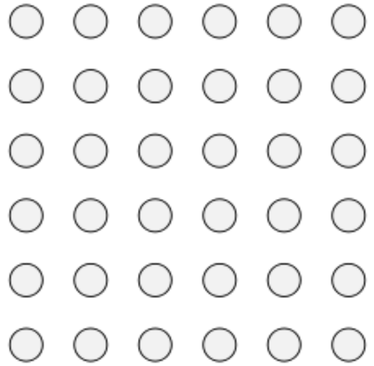


2. Law of Similarity

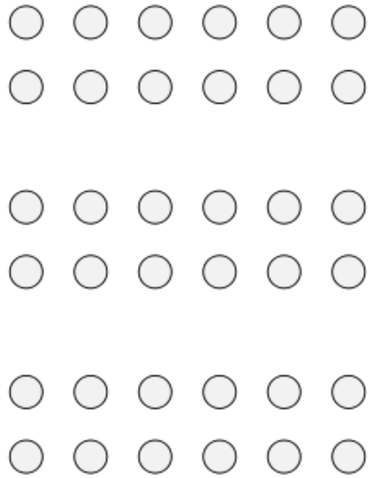
> The mind groups similar elements together



3. Law of Proximity

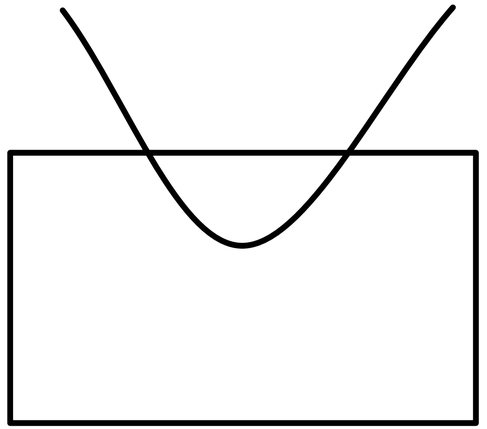


> Spatial (or temporal) proximity induces the mind to see a totality



4. Law of Continuity

> Lines follow the smoothest and simplest path.



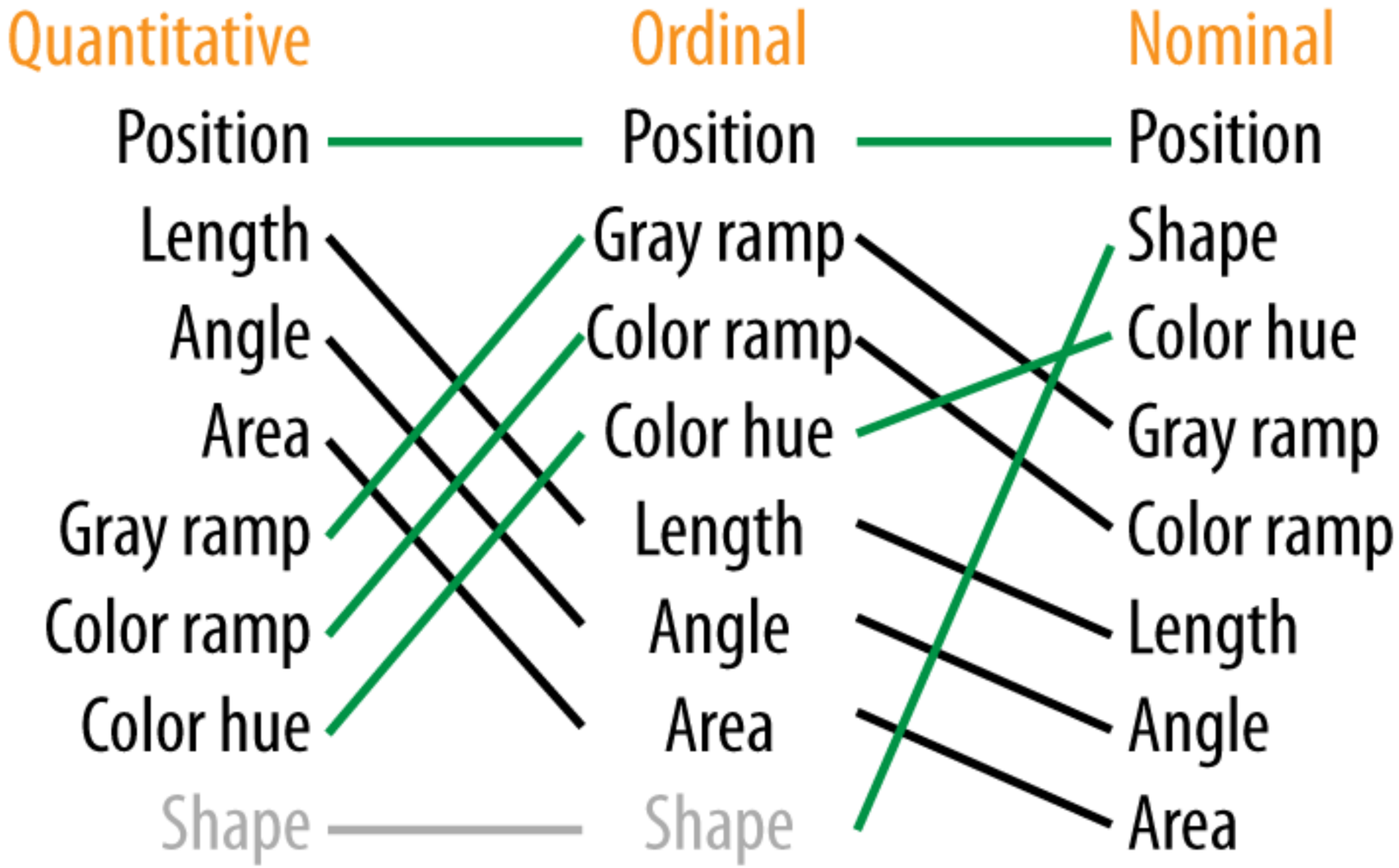
Roadmap

- Motivation
- Visual Perception
- **Information Visualization**
- Software Visualization

“The use of computer-supported interactive, ***visual representations of abstract data*** *to amplify cognition.”*

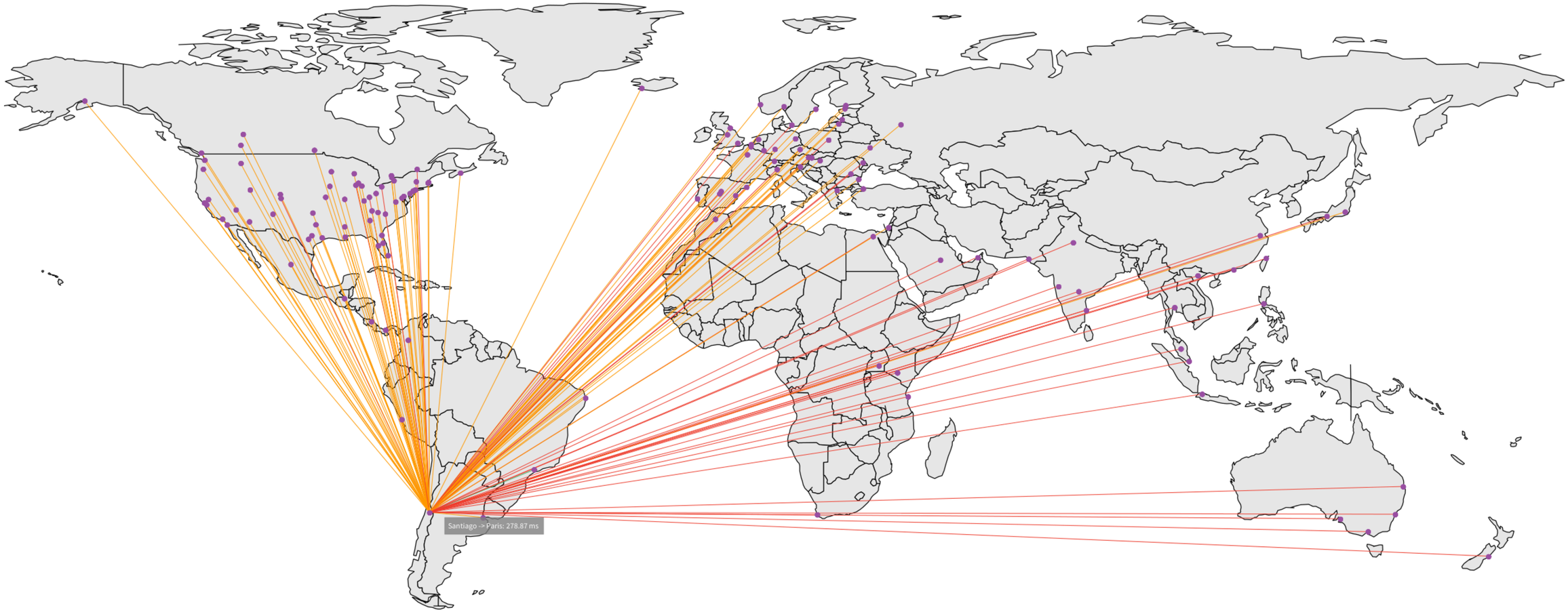
[Card et al., 1999]

Data-Driven Principles of Visualization Design

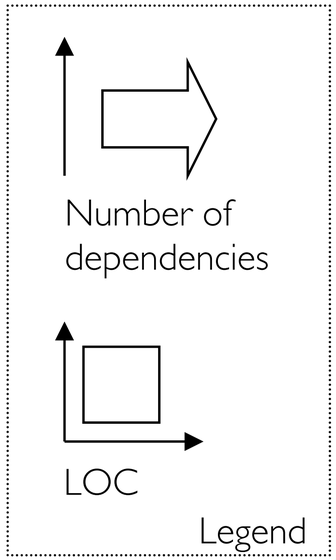


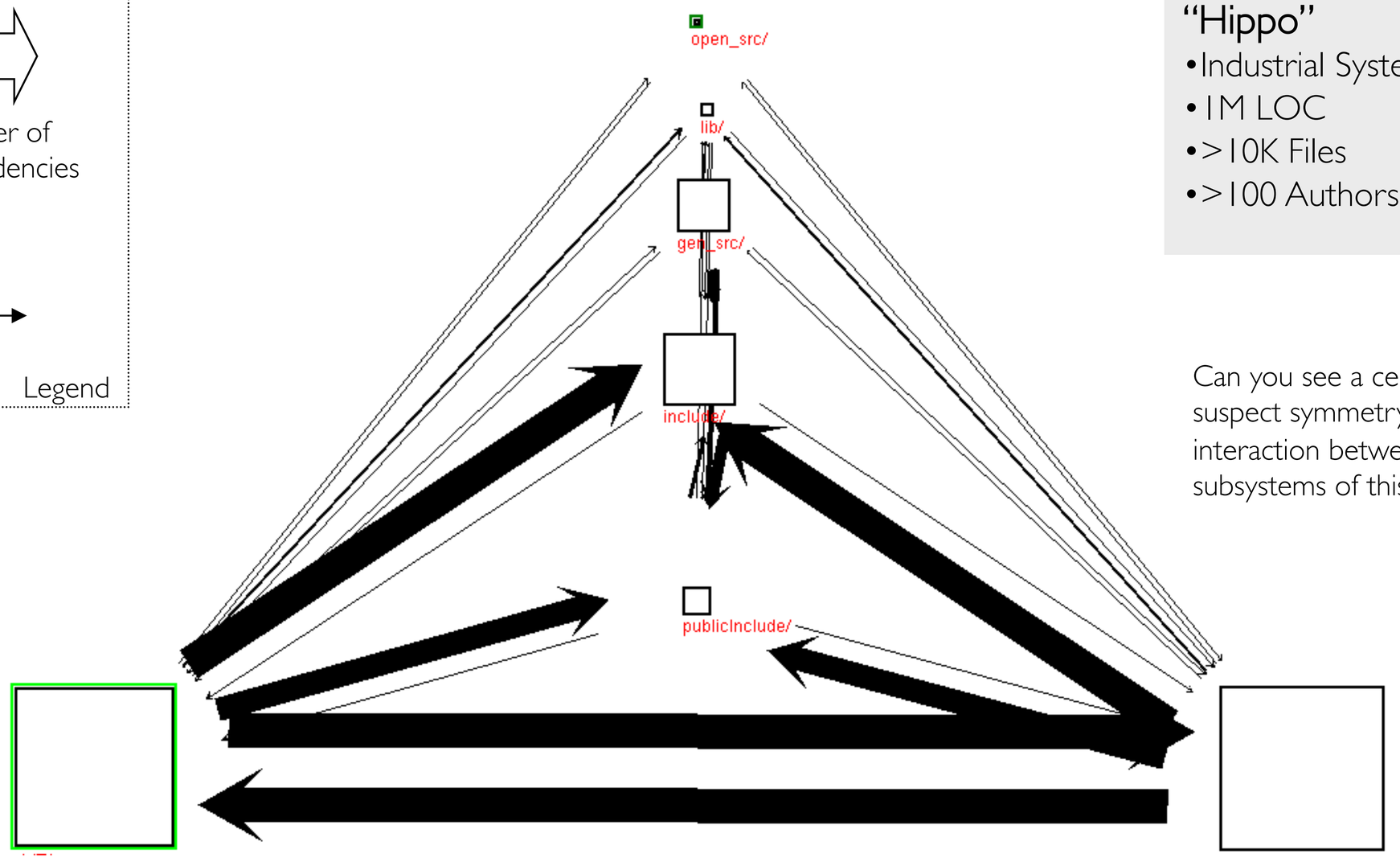
[Mackinlay, 1986]

Uncovers emergent properties and outliers



Exposes problems with the dataset

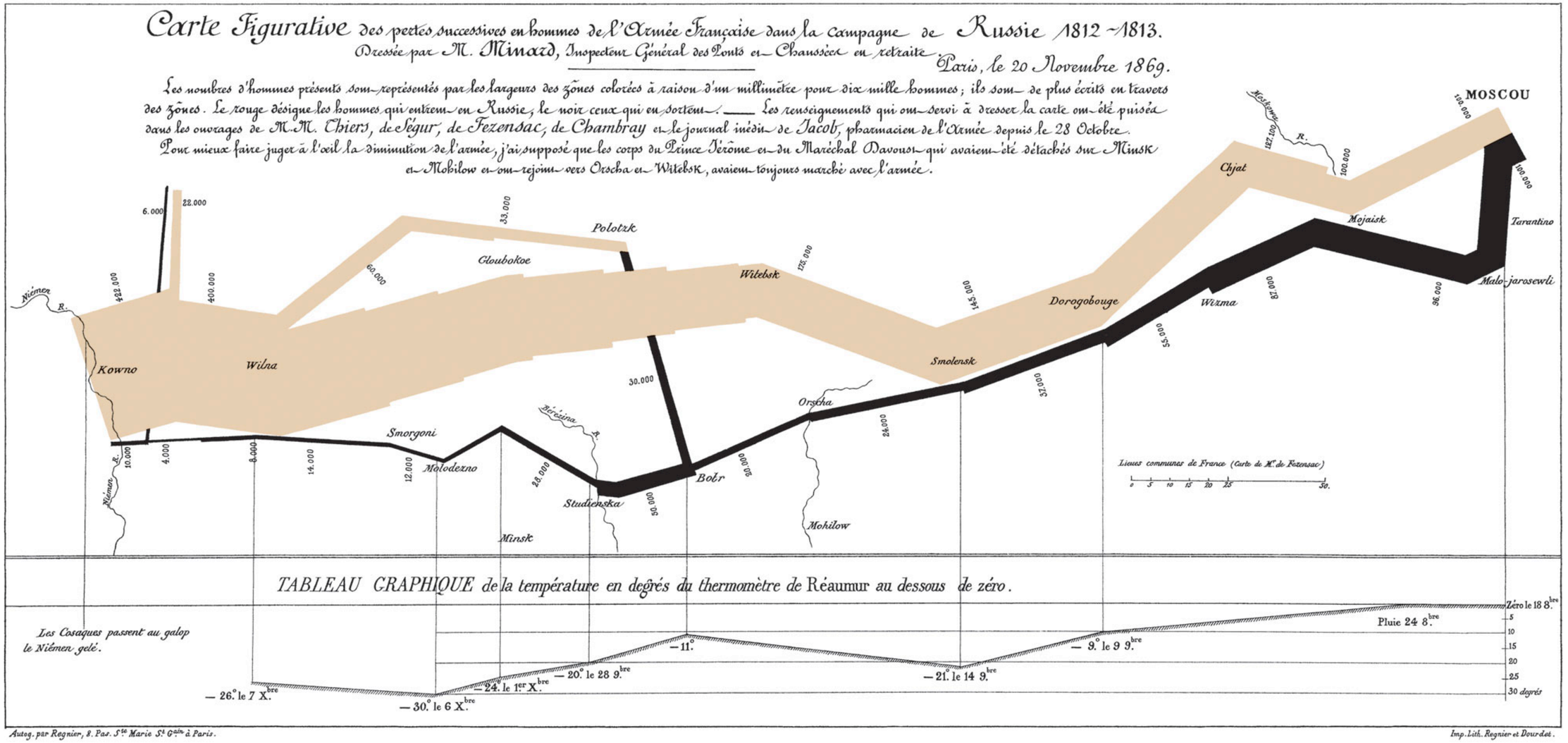




"Hippo"
- Industrial System
- 1M LOC
- >10K Files
- >100 Authors

Can you see a certain suspect symmetry in the interaction between the subsystems of this project?

Enhances communication



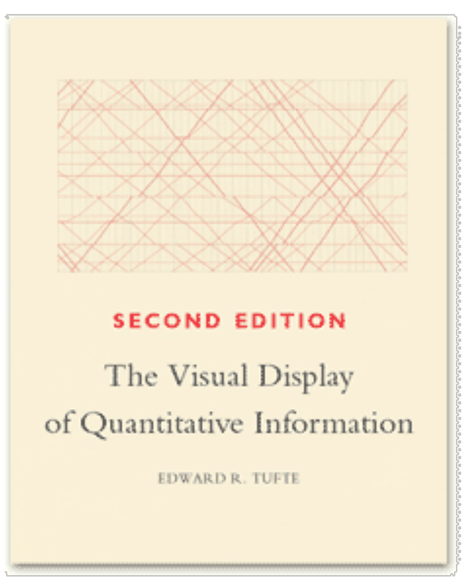
Uses of Information Visualization

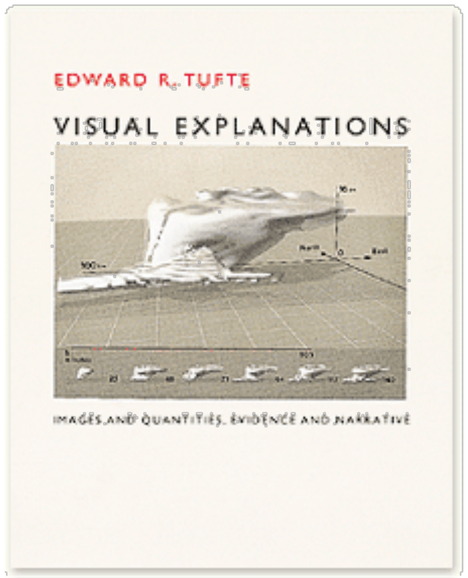
- Supports analysis
  - Uncovers emergent properties and outliers
  - Exposes problems with the data set
- Enhances communication

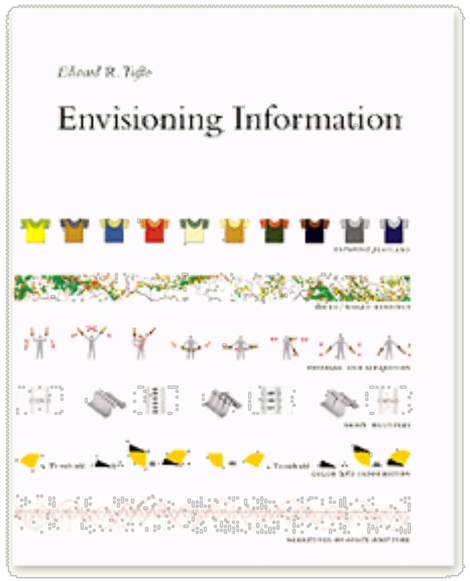
Roadmap

- Motivation
- Visual Perception
- **Information Visualization**
  - **Guidelines**
- Software Visualization

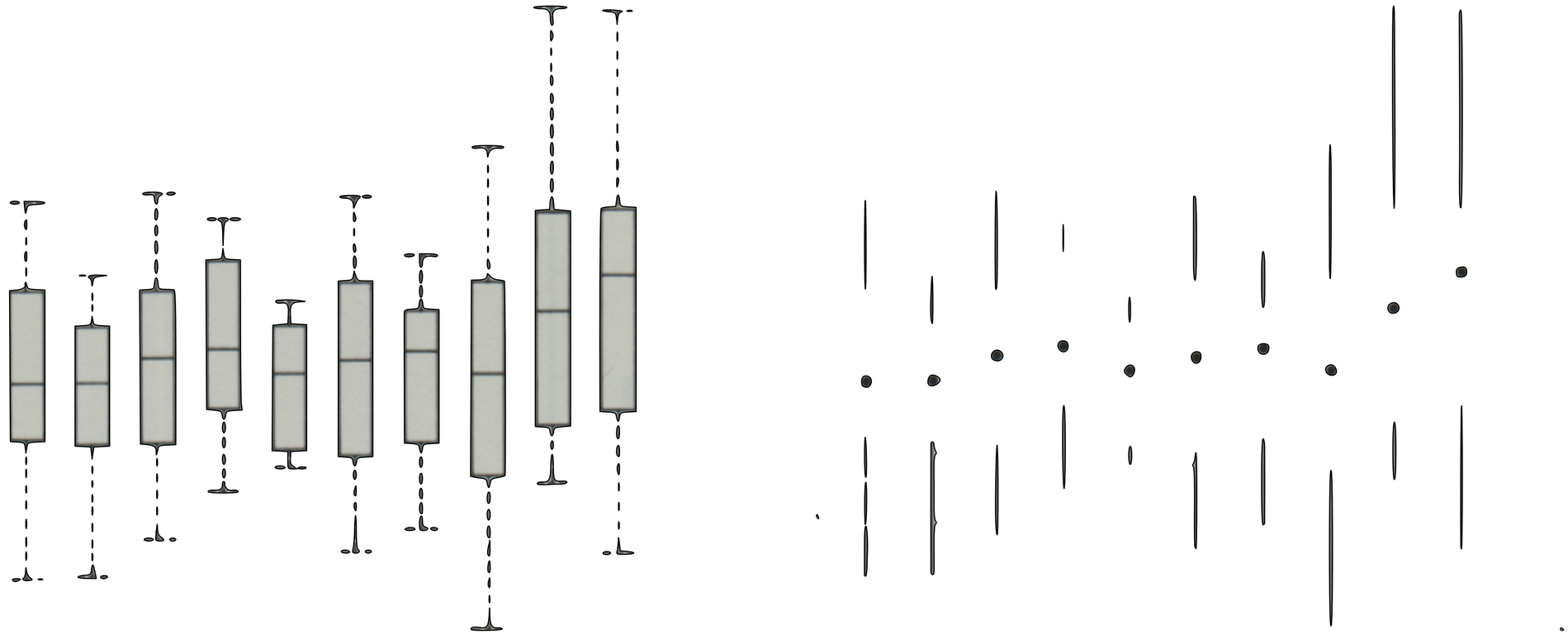
Guidelines of valuable information visualization can consider style, integrity, and design





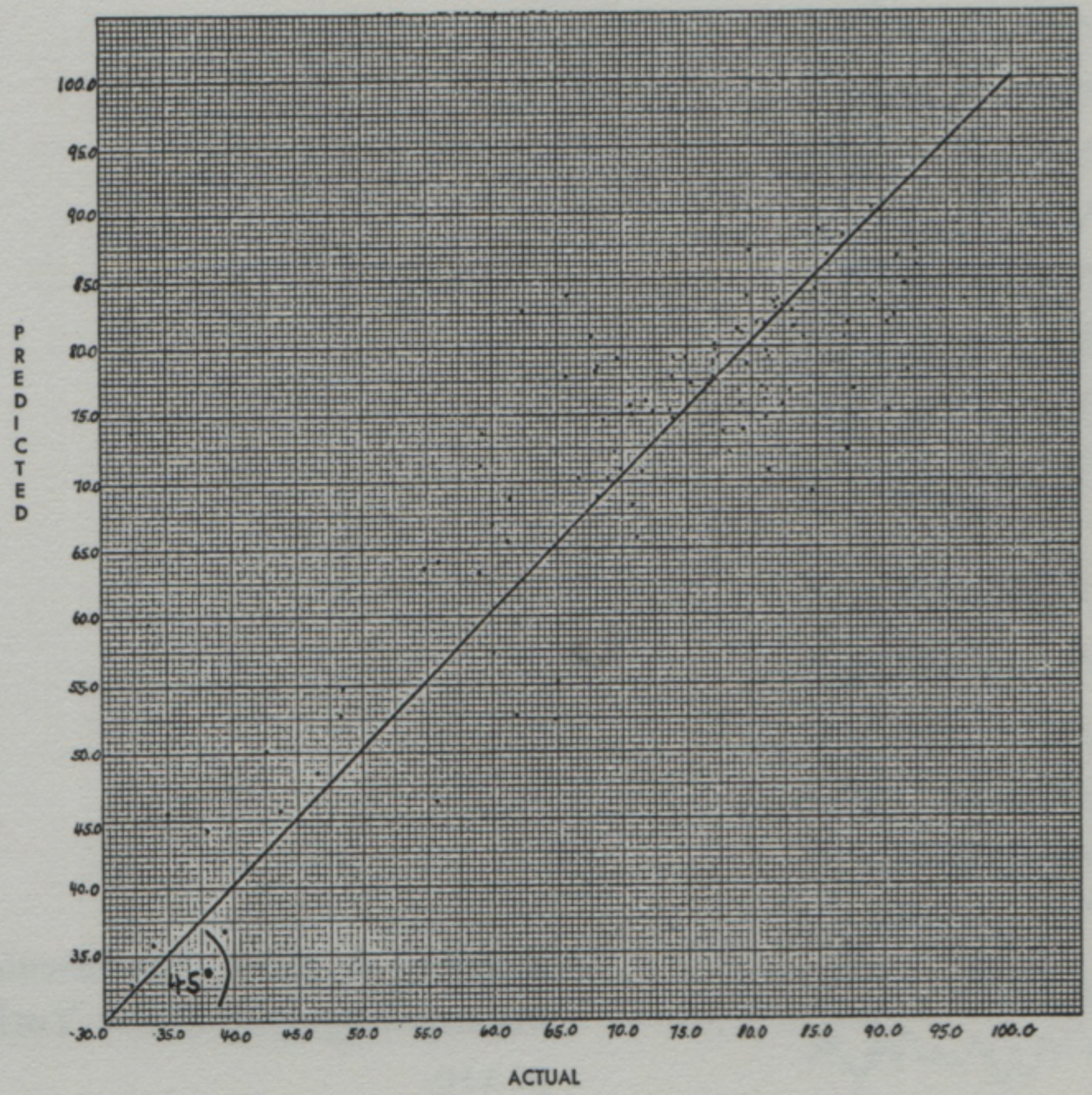


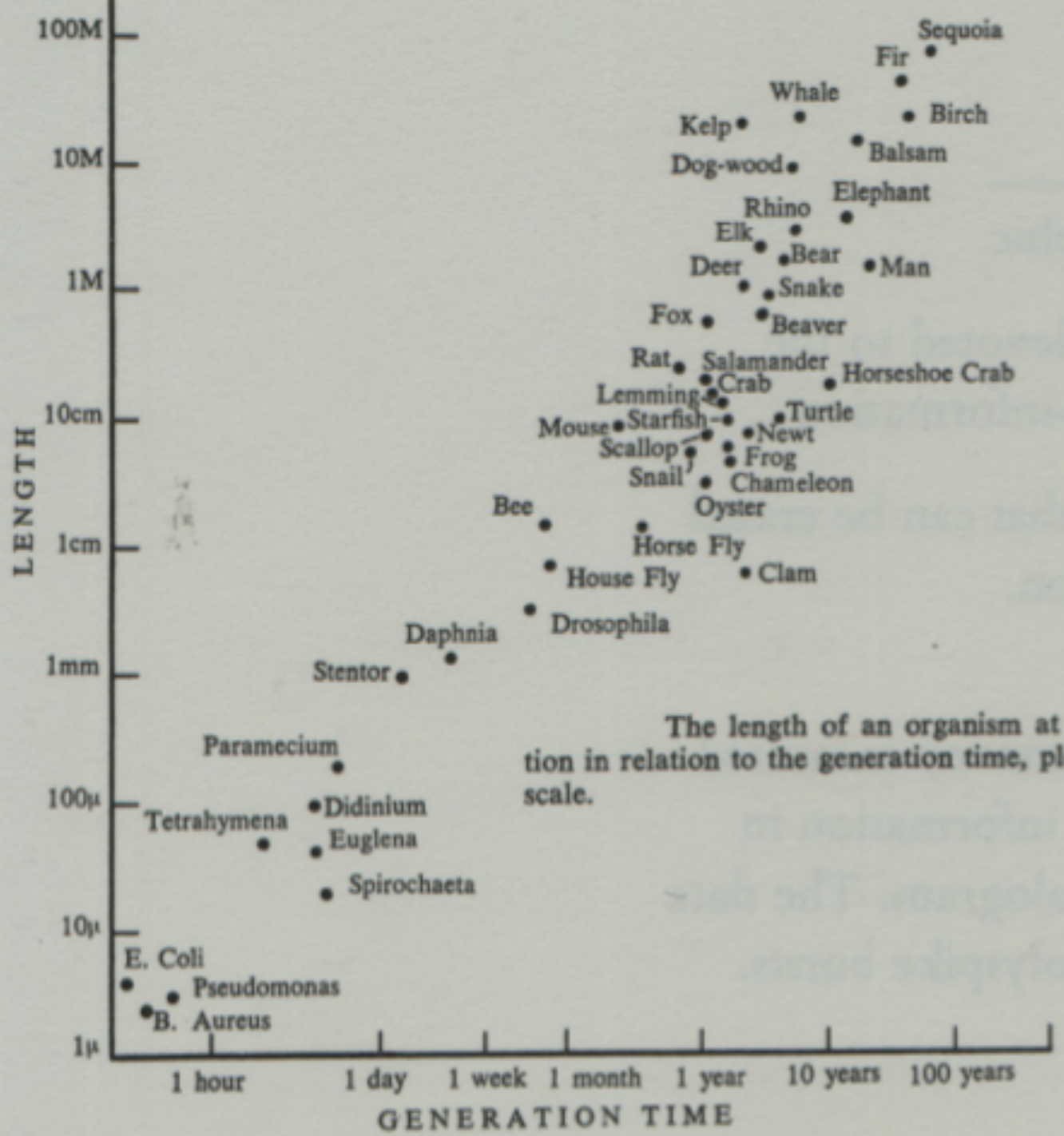
Style: Minimize Non-Data Ink



Removing ink from your graph should remove meaning from it.

Relationship of Actual Rates of Registration to Predicted Rates
(104 cities 1960).



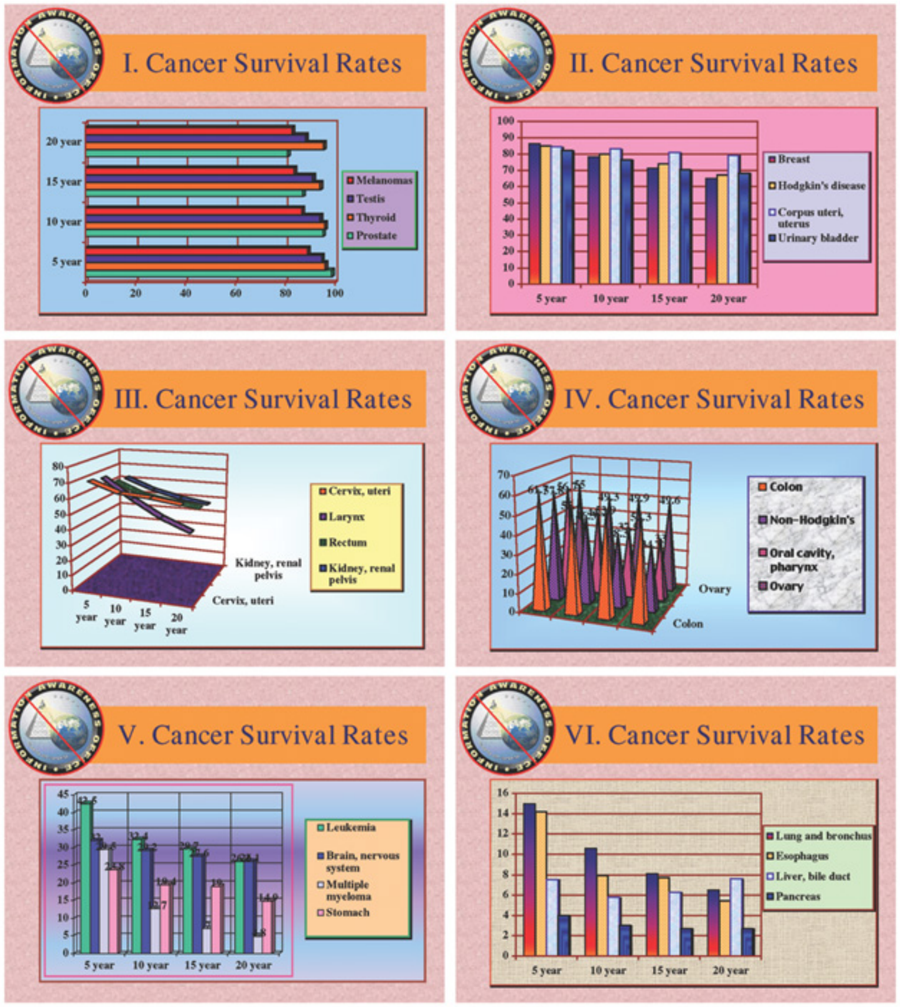


The length of an organism at the time of reproduction in relation to the generation time, plotted on a logarithmic scale.

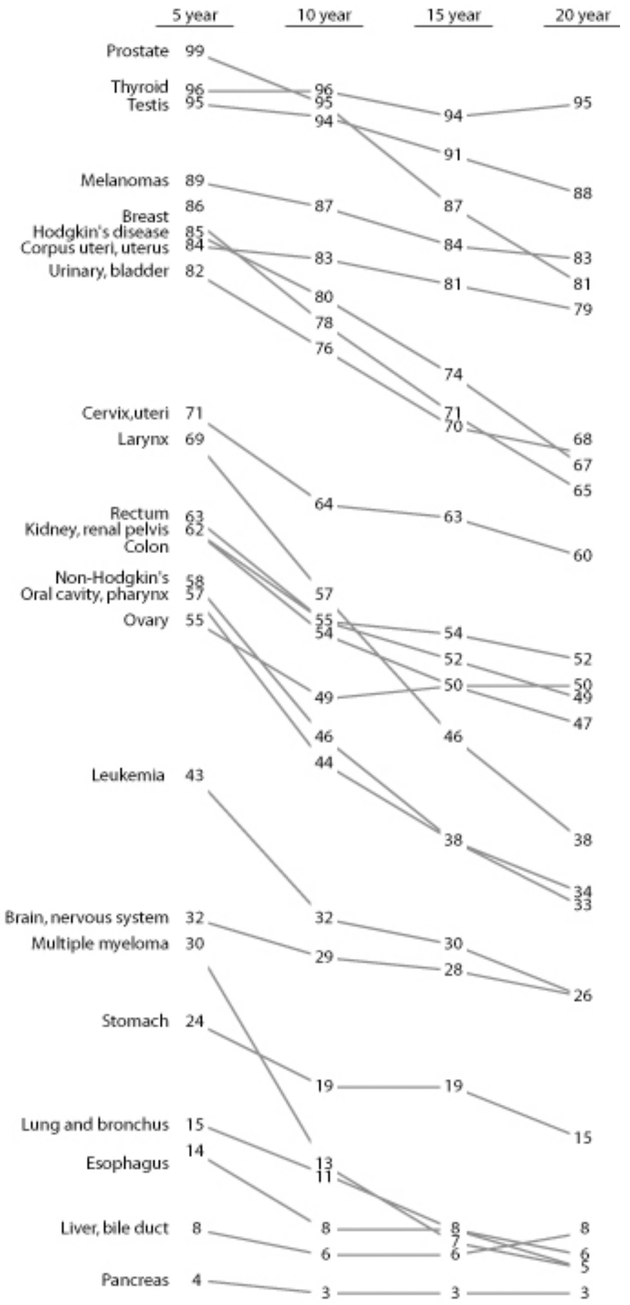
Design: Choose the appropriate representation

Estimates of relative survival rates, by cancer site

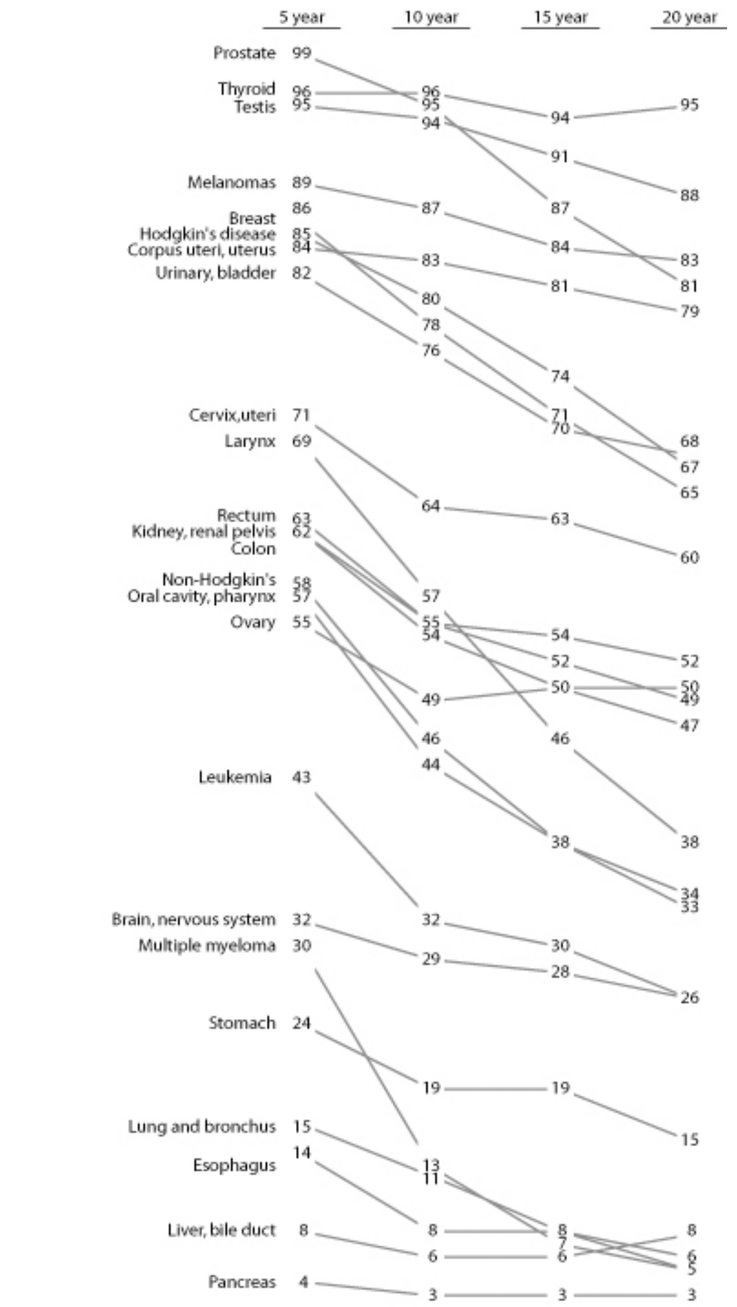
	% survival rates and standard errors							
	5 year		10 year		15 year		20 year	
Prostate	98.8	0.4	95.2	0.9	87.1	1.7	81.1	3.0
Thyroid	96.0	0.8	95.8	1.2	94.0	1.6	95.4	2.1
Testis	94.7	1.1	94.0	1.3	91.1	1.8	88.2	2.3
Melanomas	89.0	0.8	86.7	1.1	83.5	1.5	82.8	1.9
Breast	86.4	0.4	78.3	0.6	71.3	0.7	65.0	1.0
Hodgkin's disease	85.1	1.7	79.8	2.0	73.8	2.4	67.1	2.8
Corpus uteri, uterus	84.3	1.0	83.2	1.3	80.8	1.7	79.2	2.0
Urinary, bladder	82.1	1.0	76.2	1.4	70.3	1.9	67.9	2.4
Cervix, uteri	70.5	1.6	64.1	1.8	62.8	2.1	60.0	2.4
Larynx	68.8	2.1	56.7	2.5	45.8	2.8	37.8	3.1
Rectum	62.6	1.2	55.2	1.4	51.8	1.8	49.2	2.3
Kidney, renal pelvis	61.8	1.3	54.4	1.6	49.8	2.0	47.3	2.6
Colon	61.7	0.8	55.4	1.0	53.9	1.2	52.3	1.6
Non-Hodgkin's	57.8	1.0	46.3	1.2	38.3	1.4	34.3	1.7
Oral cavity, pharynx	56.7	1.3	44.2	1.4	37.5	1.6	33.0	1.8
Ovary	55.0	1.3	49.3	1.6	49.9	1.9	49.6	2.4
Leukemia	42.5	1.2	32.4	1.3	29.7	1.5	26.2	1.7
Brain, nervous system	32.0	1.4	29.2	1.5	27.6	1.6	26.1	1.9
Multiple myeloma	29.5	1.6	12.7	1.5	7.0	1.3	4.8	1.5
Stomach	23.8	1.3	19.4	1.4	19.0	1.7	14.9	1.9
Lung and bronchus	15.0	0.4	10.6	0.4	8.1	0.4	6.5	0.4
Esophagus	14.2	1.4	7.9	1.3	7.7	1.6	5.4	2.0
Liver, bile duct	7.5	1.1	5.8	1.2	6.3	1.5	7.6	2.0
Pancreas	4.0	0.5	3.0	1.5	2.7	0.6	2.7	0.8



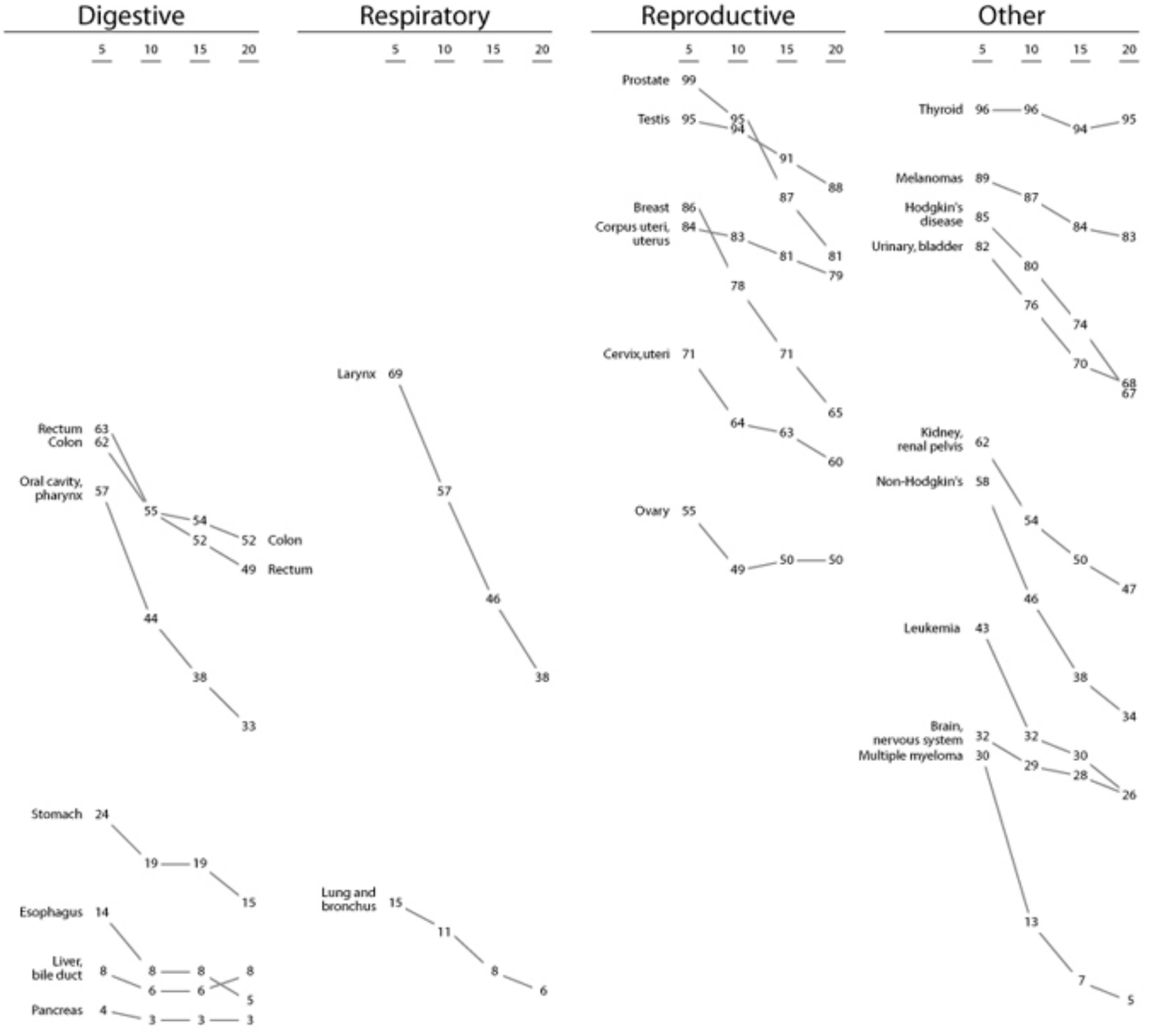
Estimates of relative survival rates, by cancer site



Estimates of relative survival rates, by cancer site

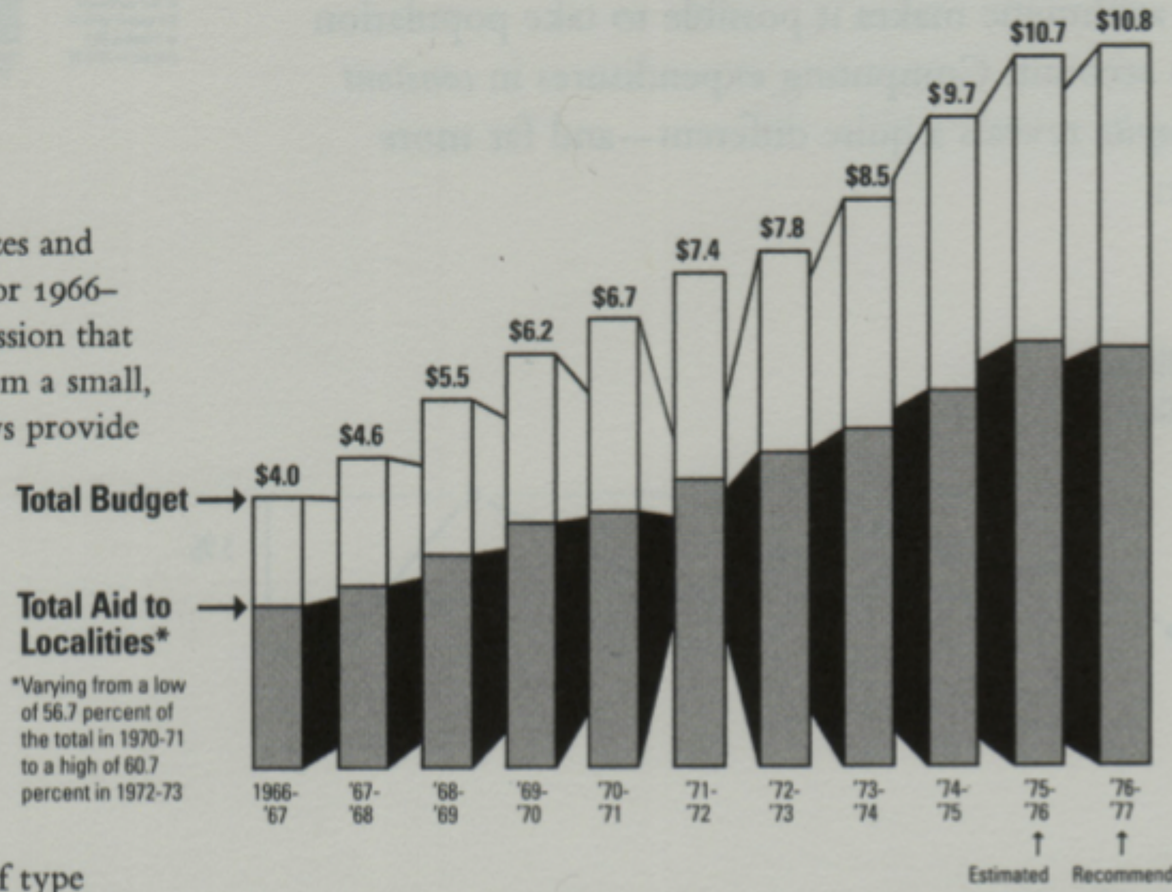


Estimates of relative survival rates, by cancer site



Integrity: Present only the data

This cluster of type emphasizes and stretches out the low value for 1966–1967, encouraging the impression that recent years have shot up from a small, stable base. Horizontal arrows provide similar emphasis.

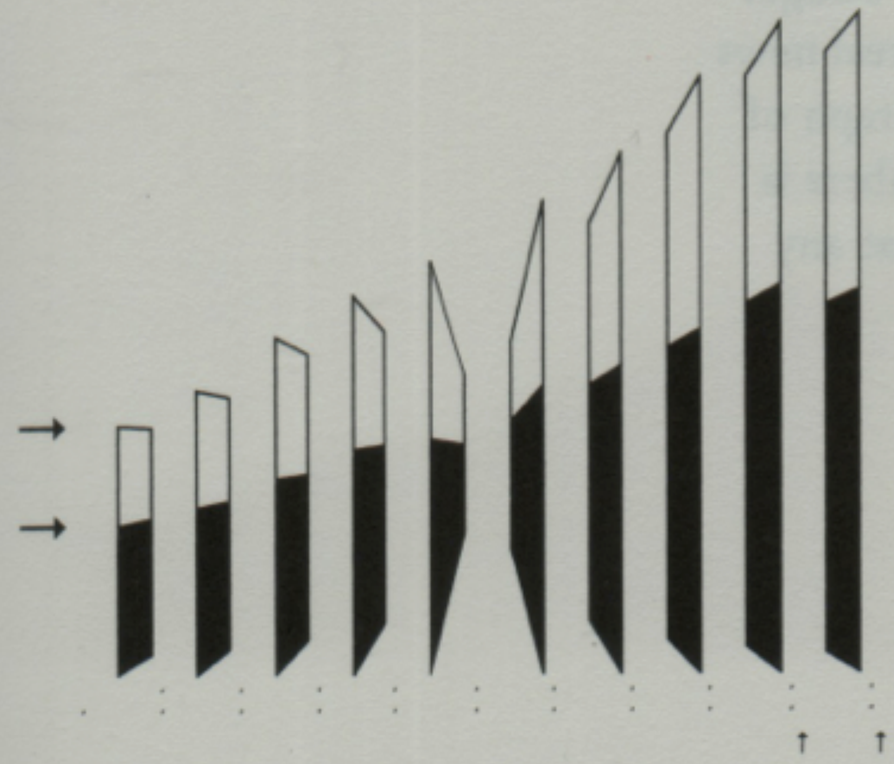


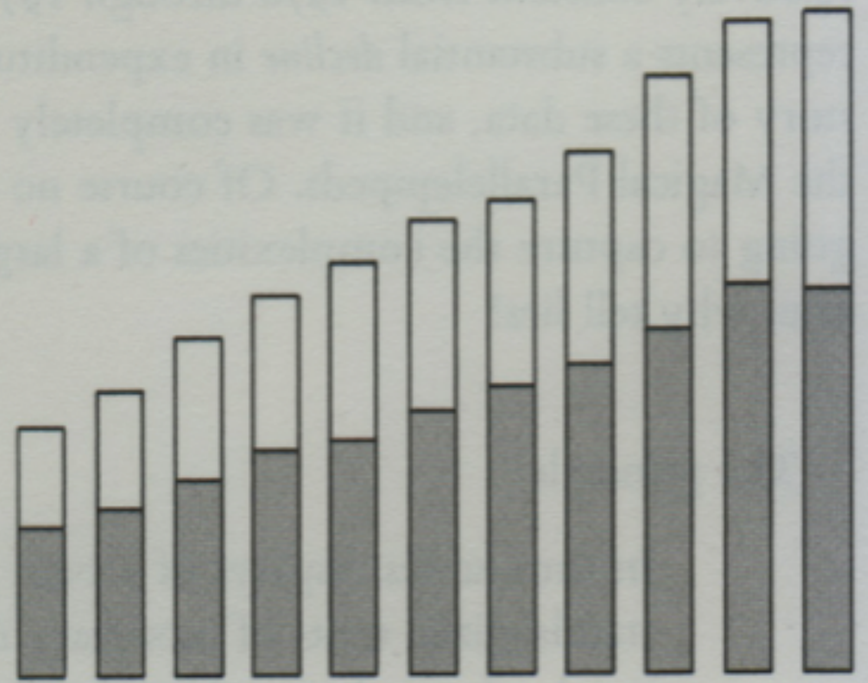
This squeezed-down block of type contributes to an image of small, squeezed-down budgets back in the good old days.

Arrows pointing straight up emphasize recent growth. Compare with horizontal arrows at left.

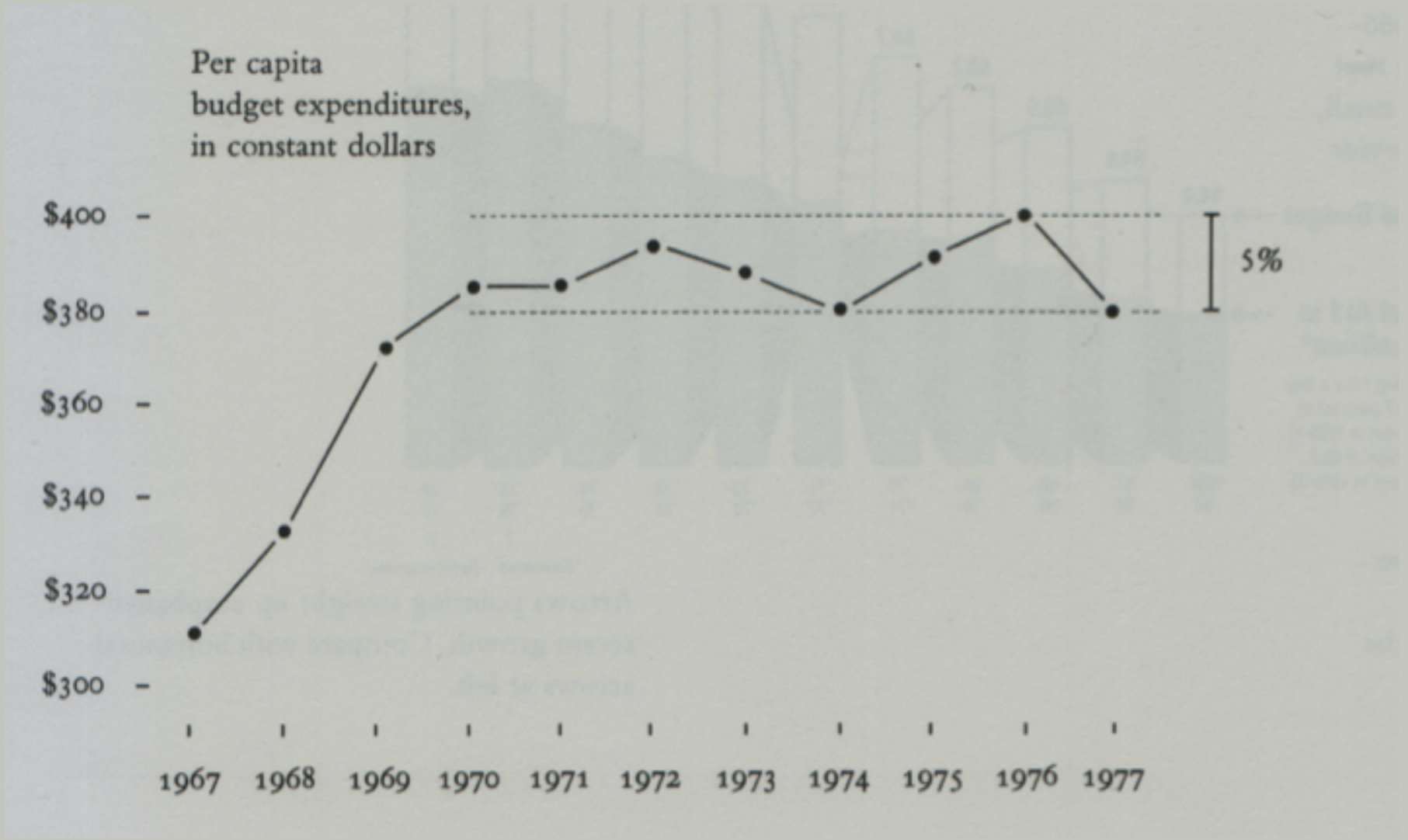
Improvement 1: Eliminate Chart Junk

Leaving behind the distortion in the chartjunk heap at the left yields a calmer view:





Improvement 2: Adjust the underlying information...



Roadmap

- Motivation
- Visual Perception
- Information Visualization
- **Software Visualization**

“SV is the use of interactive computer graphics to enhance the interface between users and their programs.”

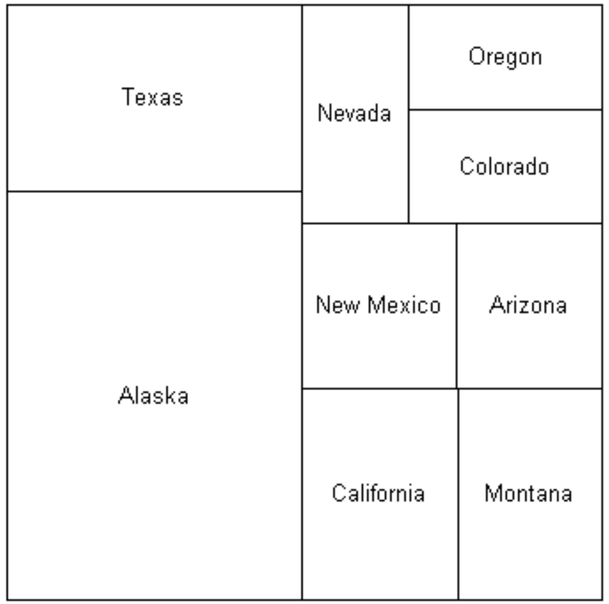
[Price, 1992]

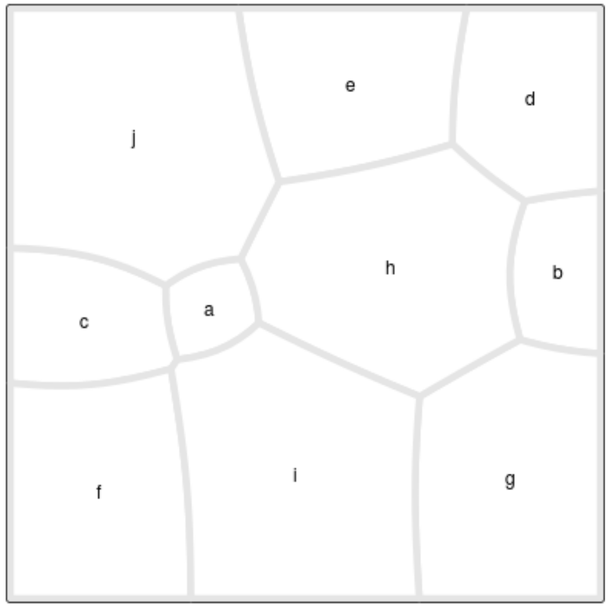
Roadmap

- Motivation
- Visual Perception
- Information Visualization
- **Software Visualization**
  - **Structure**
  - Evolution
  - Behavior

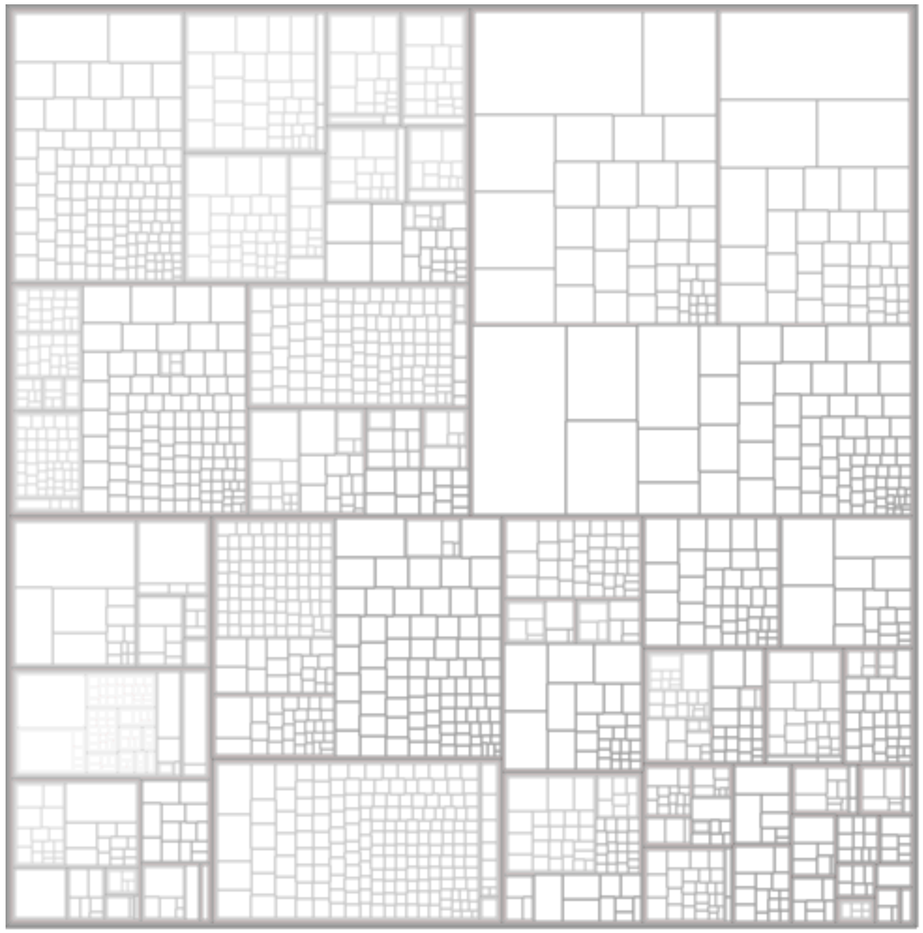
Space Filling Techniques

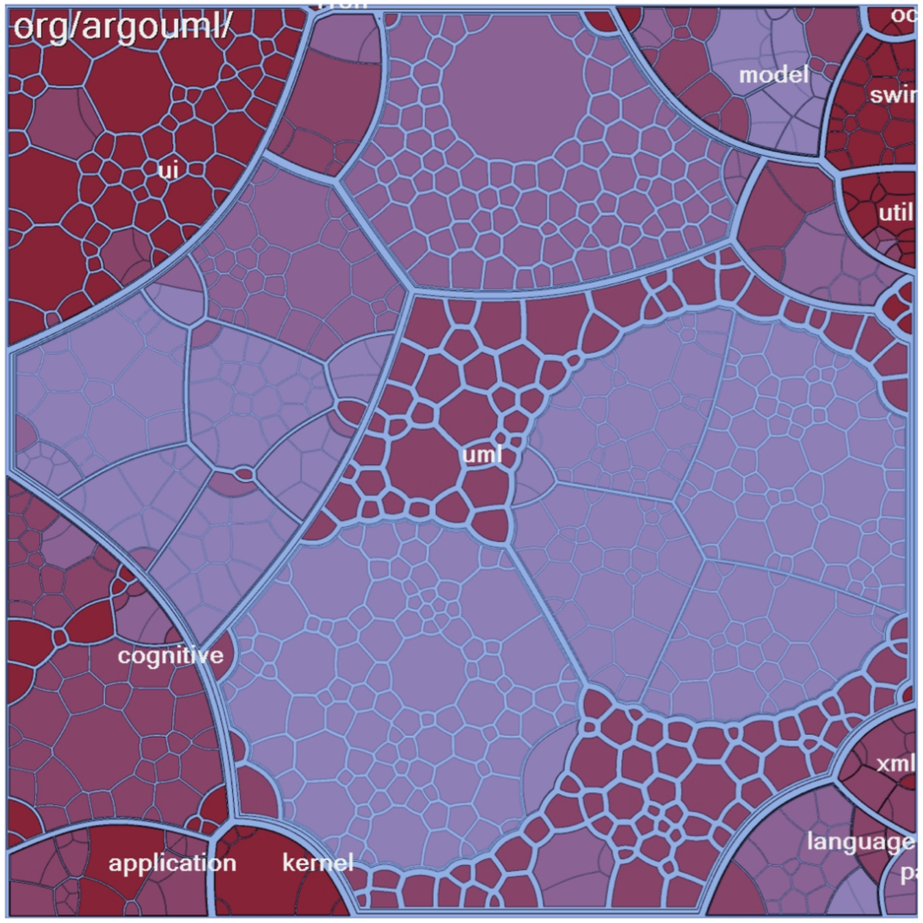
- Use of pre-attentive processing features of
  - Locality
  - Size
- Techniques
  - Treemaps
  - Voronoi diagrams





Providing an overview of size distribution

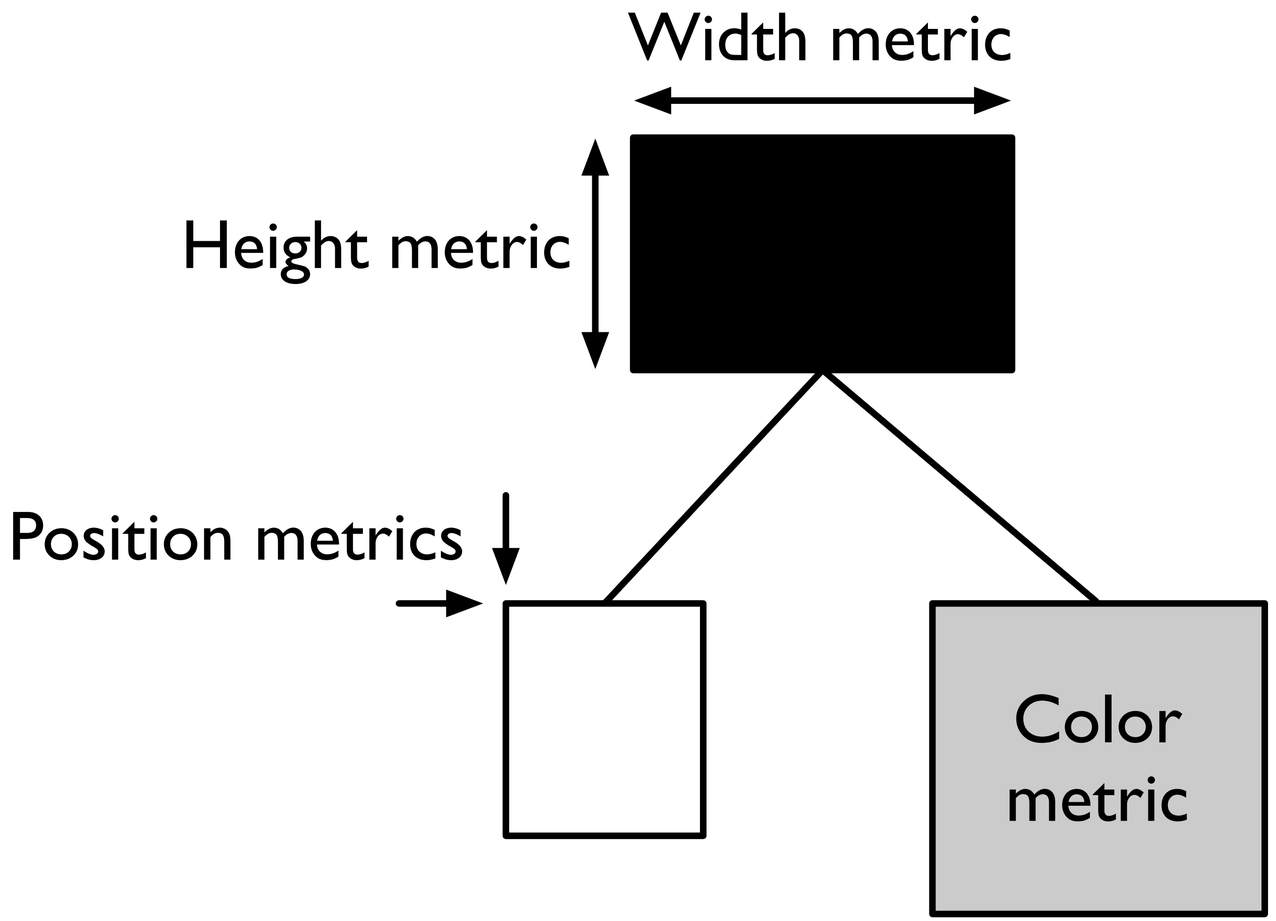




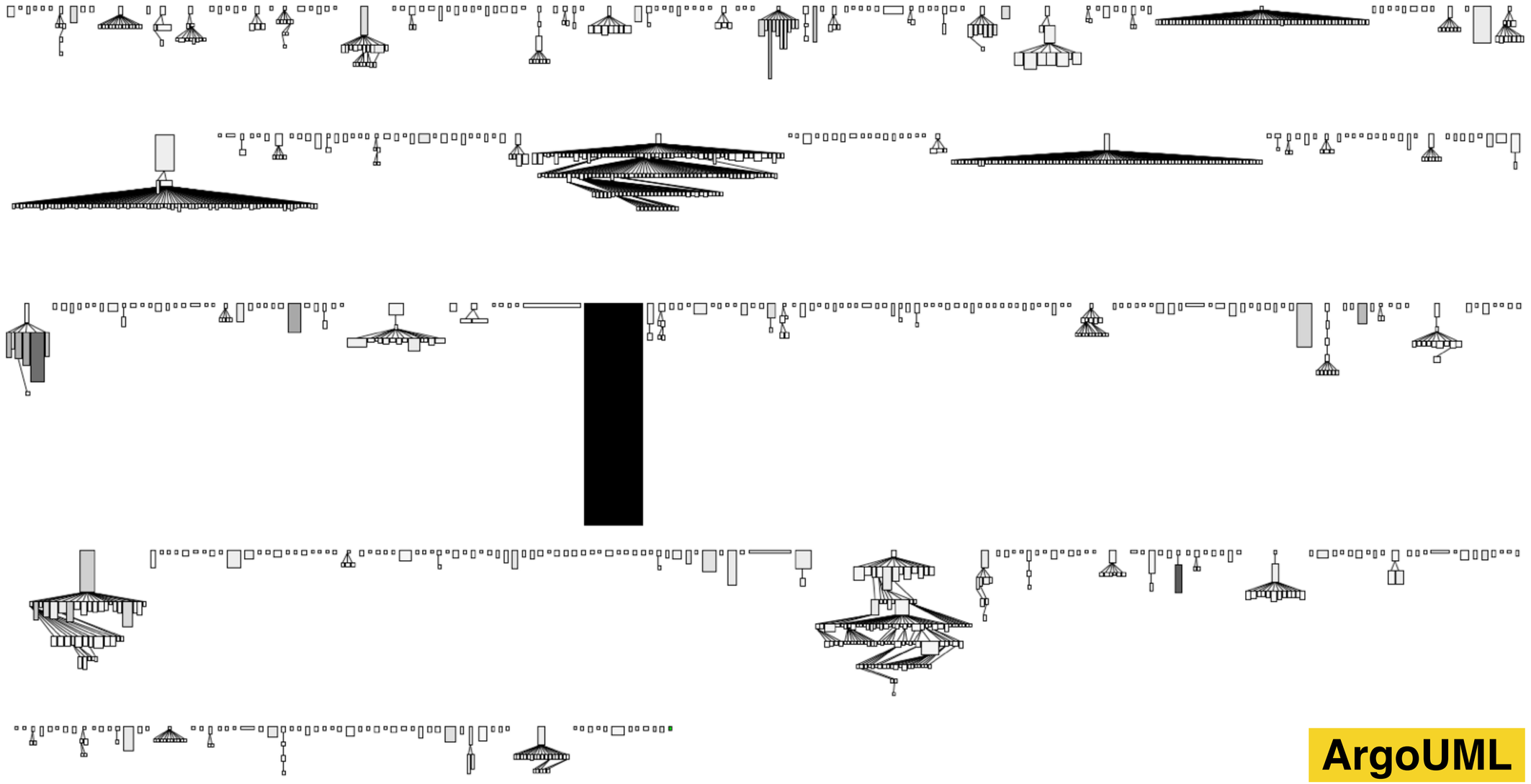
ArgoUML

Polymetric Views

- Use of pre-attentive processing features
  - Size
  - Color
  - Connectedness
- Available tools
  - Roassal, Mondrian, XRay



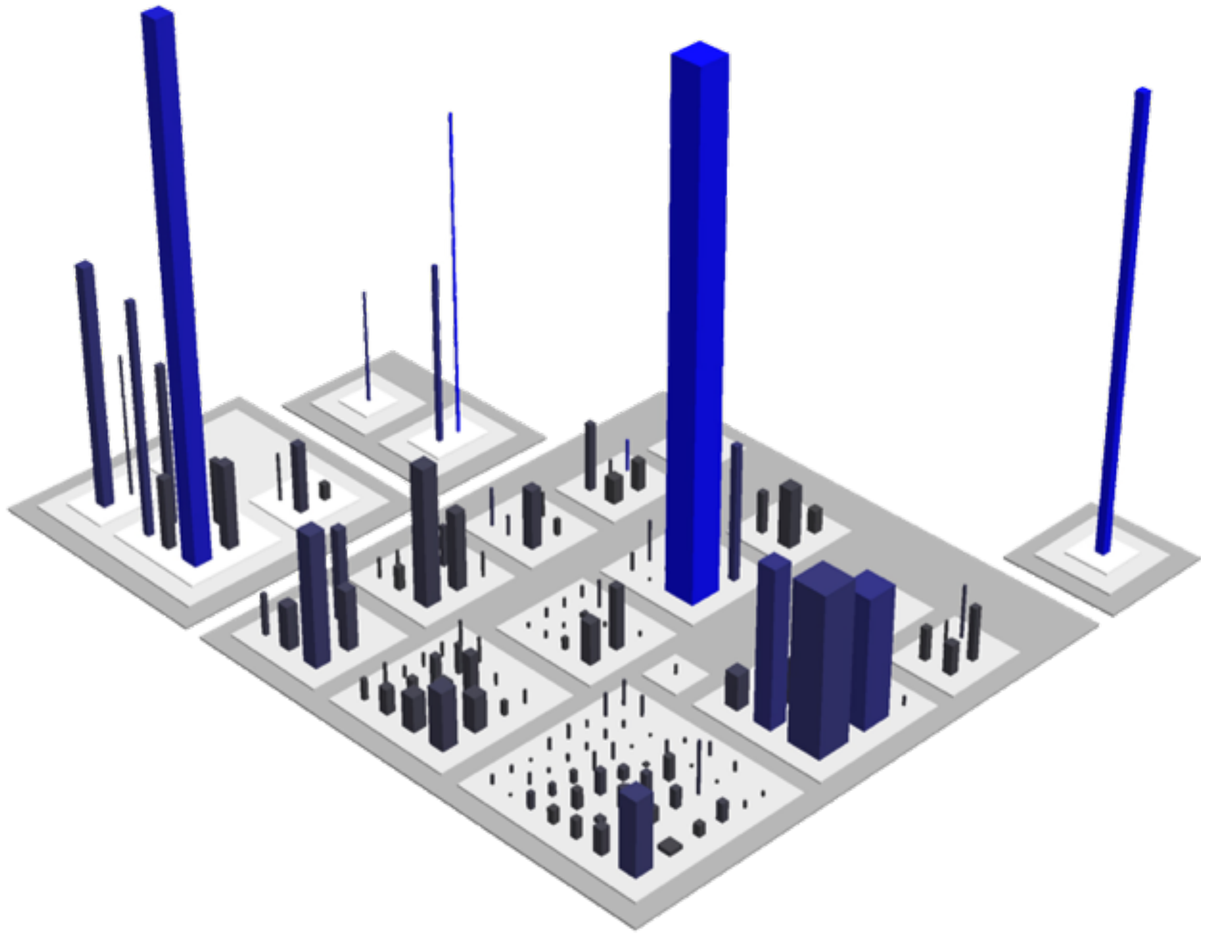
Providing an overview of inheritance



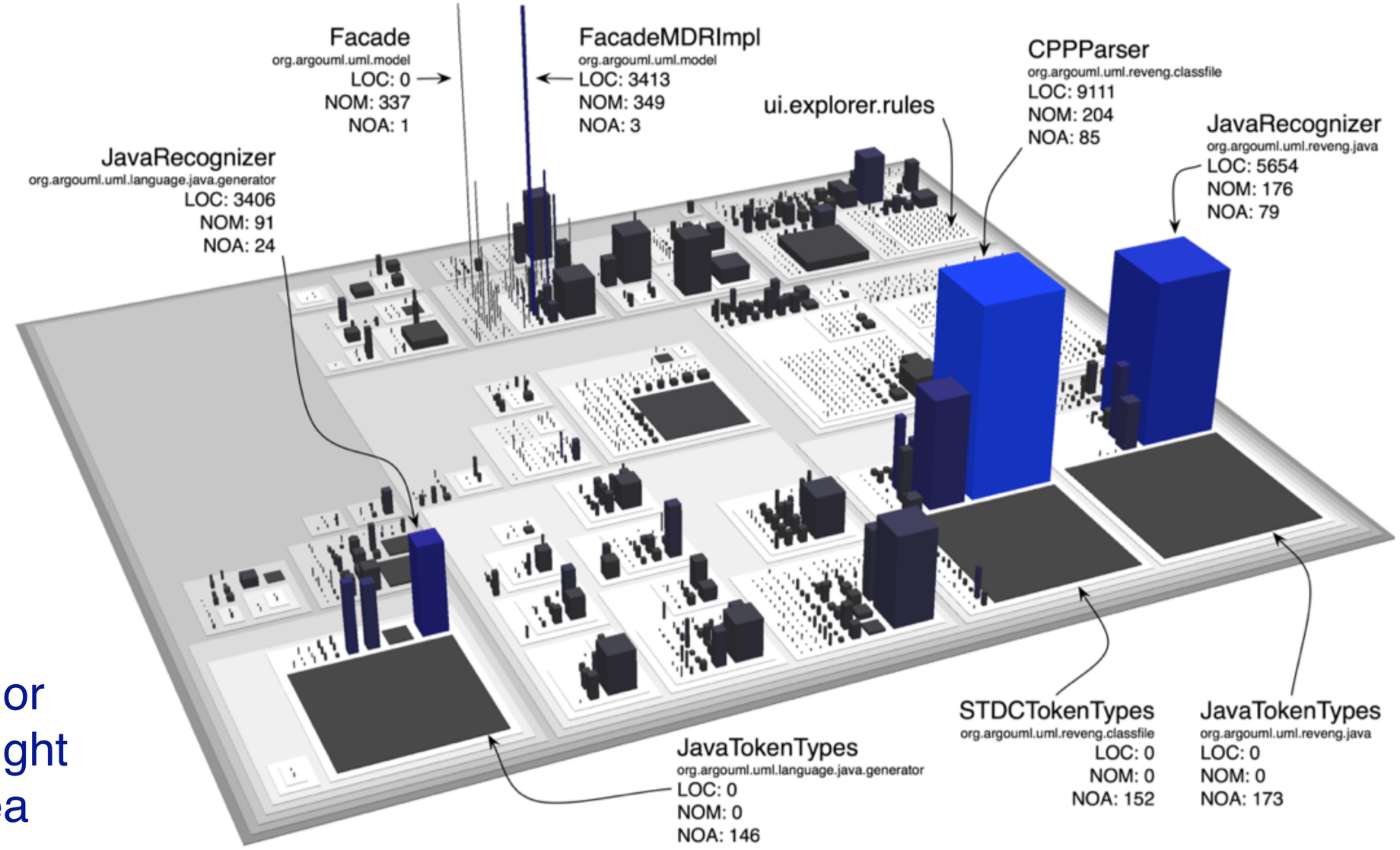
ArgoUML

Metaphor: City

- Use of pre-attentive processing features of
  - Size
  - Color
  - 3D spatial locality
- Available tools
  - CodeCity
  - CityVR

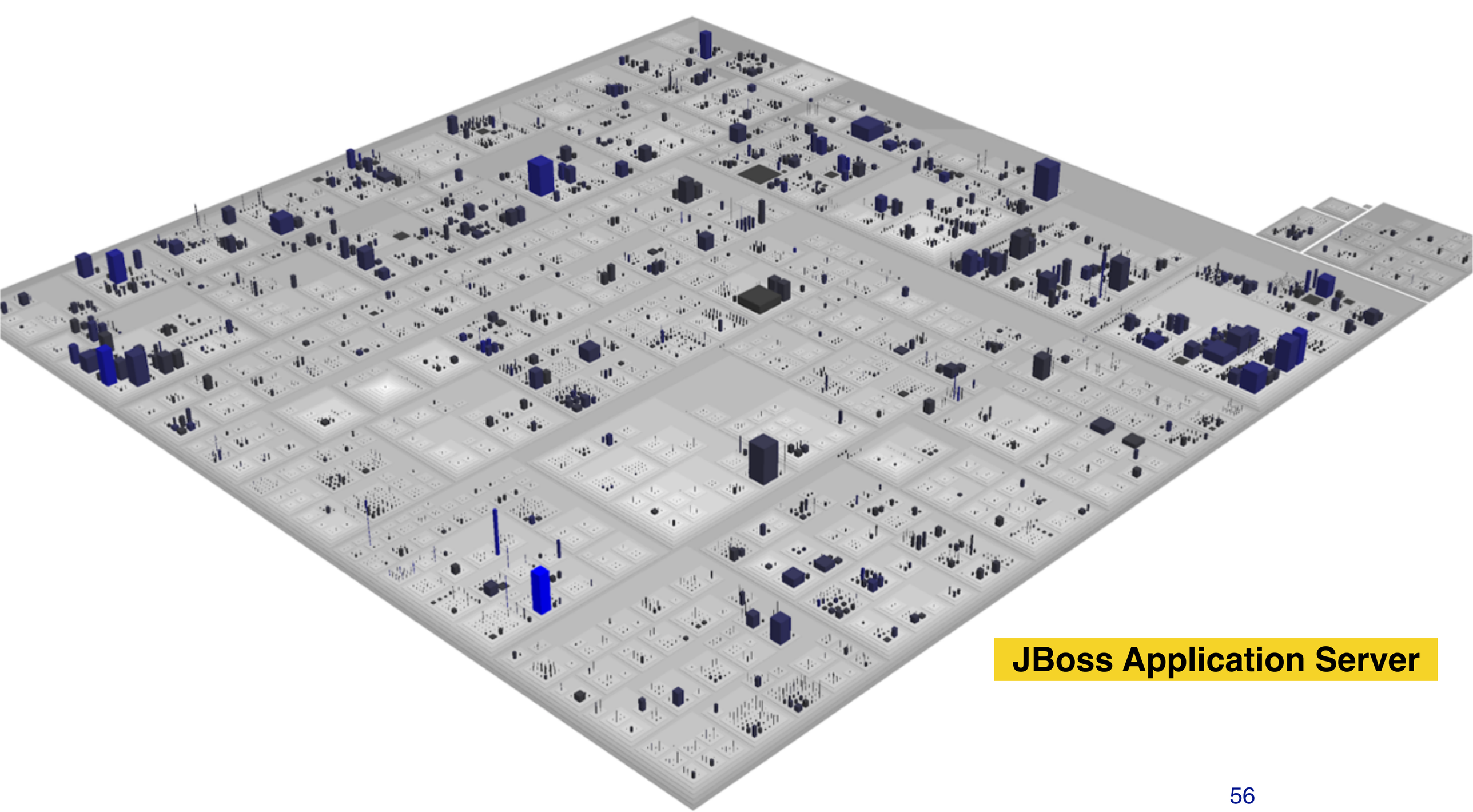


Detecting Outliers



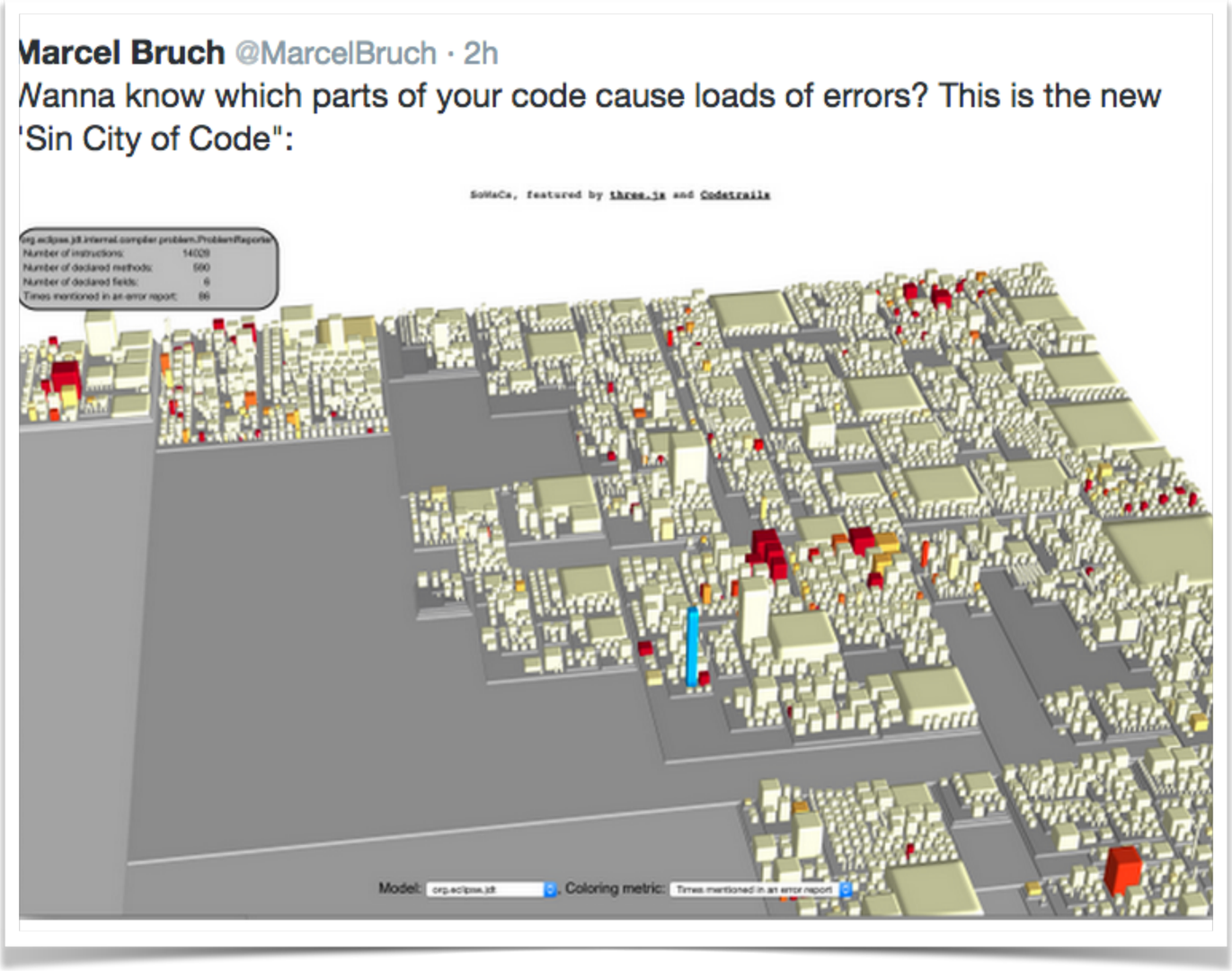
LOC -> Color
NOM -> Height
NOA -> Area

ArgoUML



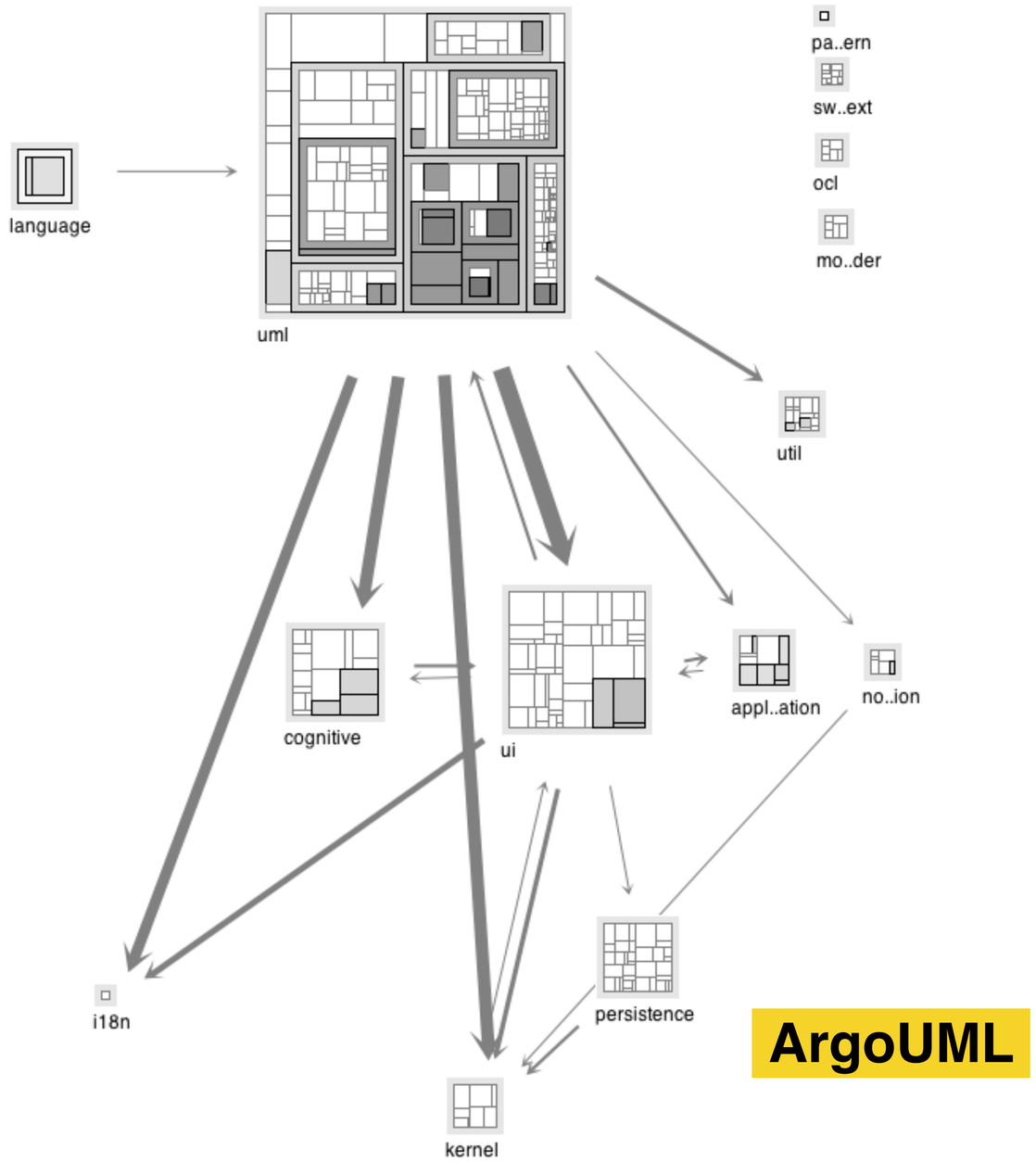
JBoss Application Server

Communicating the locality of problems



Hierarchical Visualization

- Use of pre-attentive processing features of
  - Size
  - Spatial locality
  - Connectedness
  - Color
- Available tools
  - Softwarenaut, Rigi, Shrimp, etc.



ArgoUML

Structure — Summary

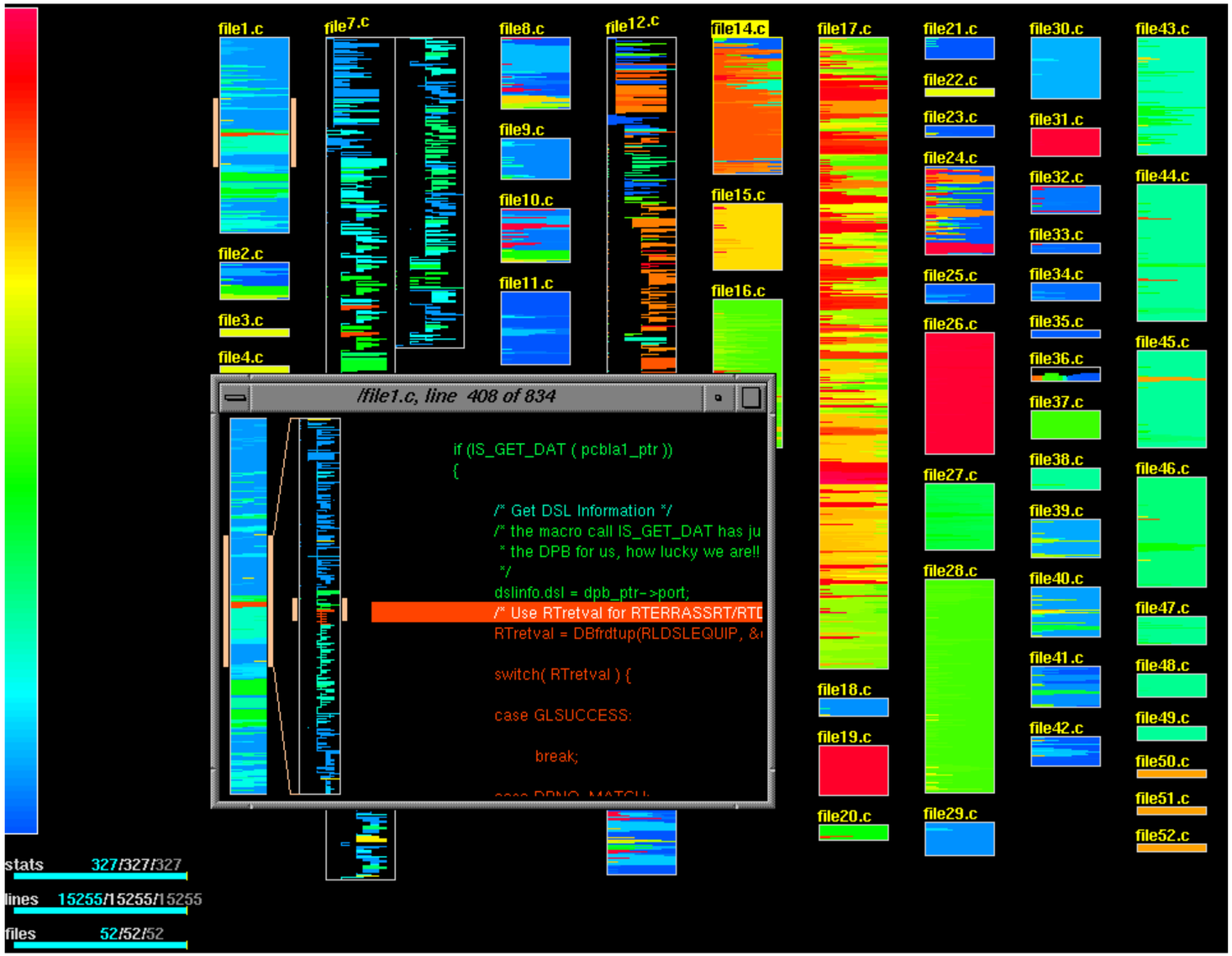
- Visualized software aspects
  - Inheritance
  - Containment
  - Dependencies
- Visualization techniques
  - Space filling techniques
  - Polymetric Views
  - Metaphors
  - Hierarchical Visualization

Roadmap

- Visual Perception
- Information Visualization
- **Software Visualization**
  - Structure
  - **Evolution**
  - Behavior

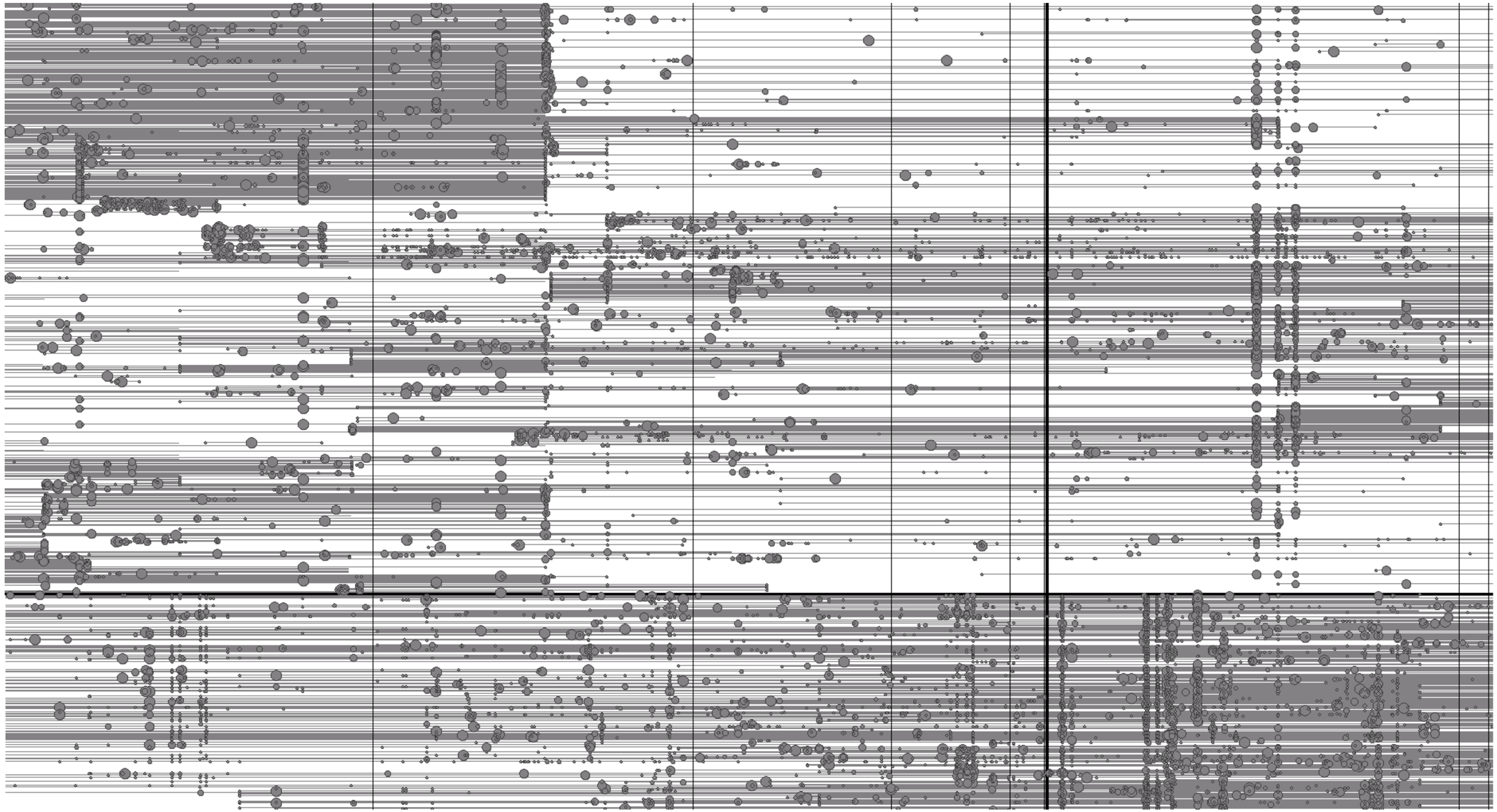
CODE_SWARM

SeeSoft mapped time on color to visualize churn

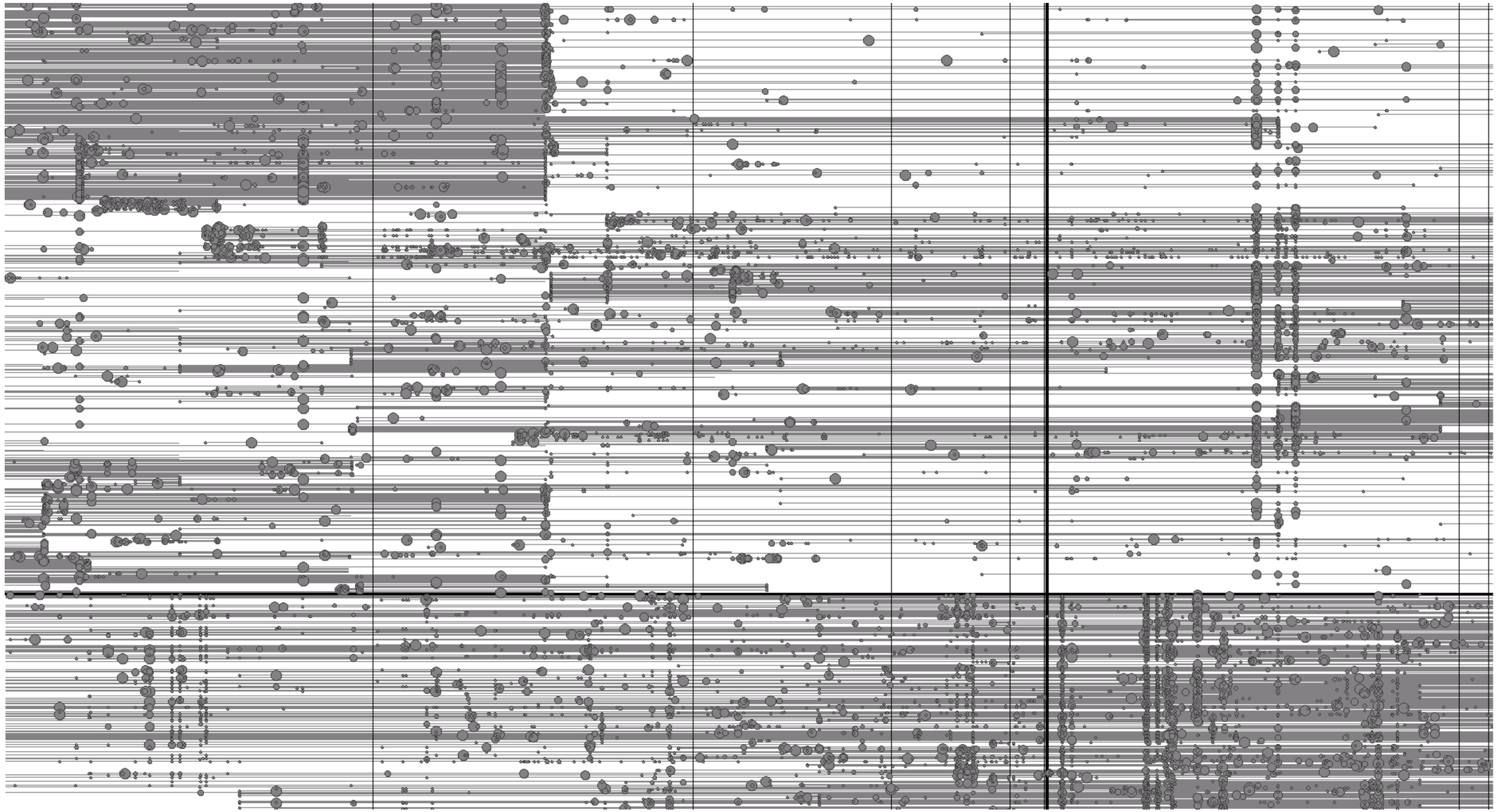


The Seesoft system maps **each line of code into a thin row**. The color of each row indicates a statistic of interest, e.g., **red rows are those most recently changed**, and **blue are those least recently changed**

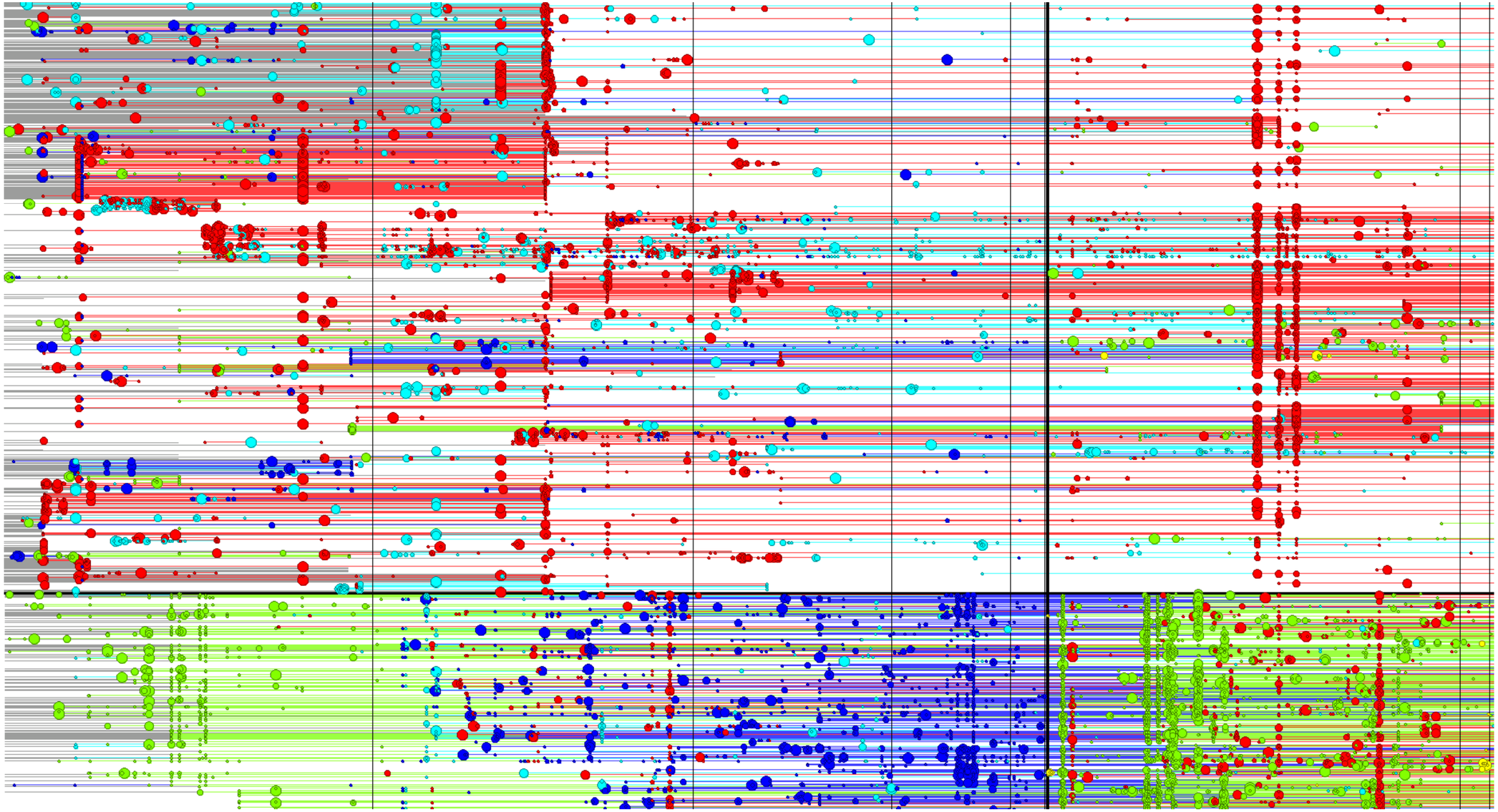
Mapping evolution on space (the x-axis)



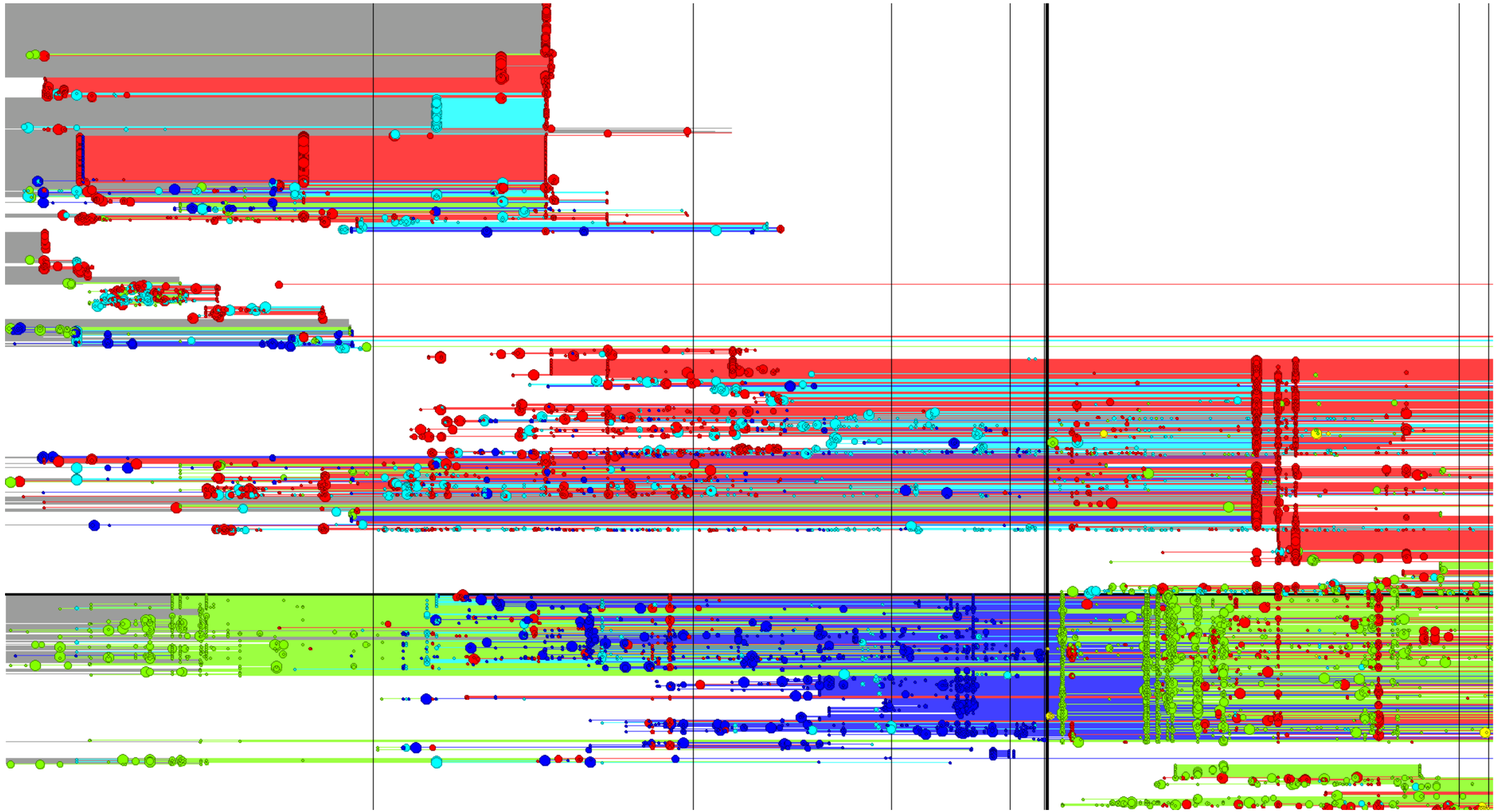
Y-axis represents individual files sorted alphabetically



Map authors on colors and kill alphabetical order

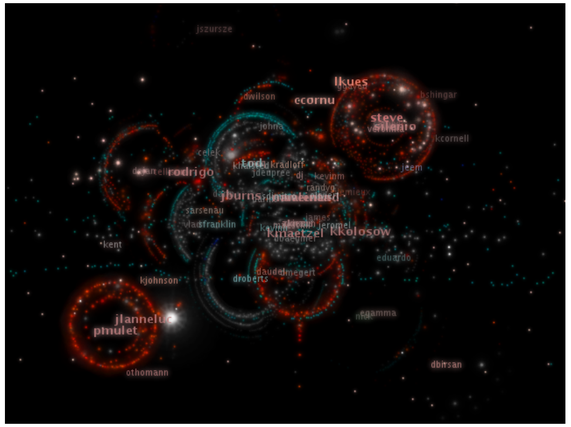


Alphabetical order must die (J. Nielsen)

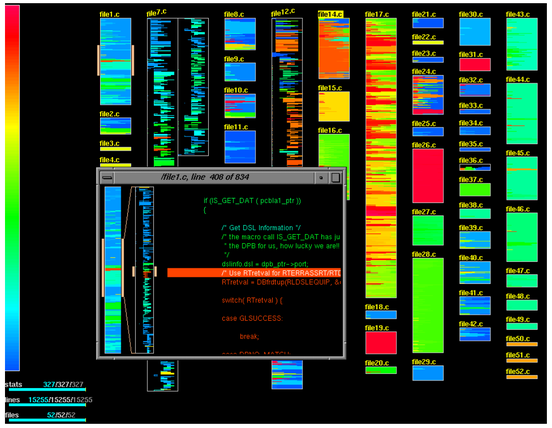


Girba etal,

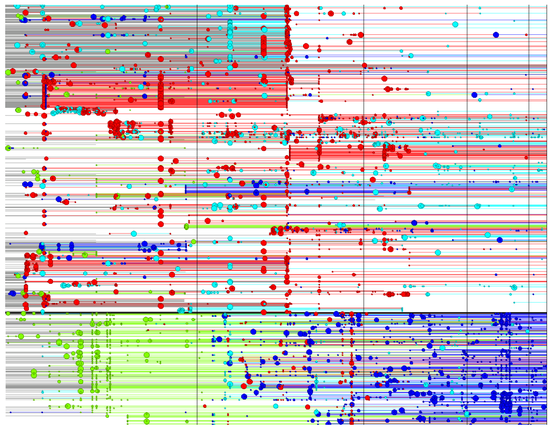
System evolution can be mapped on



time



color

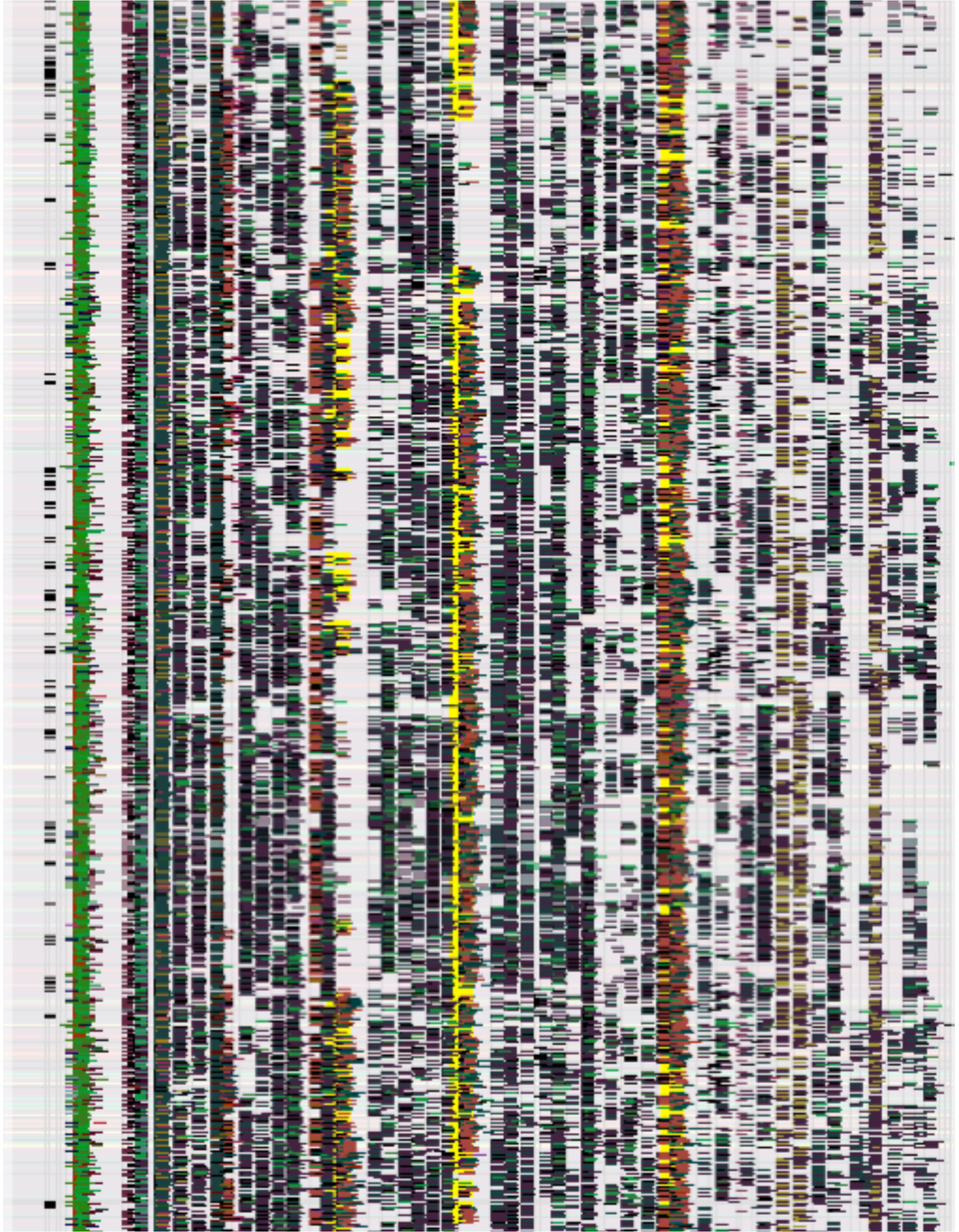


space

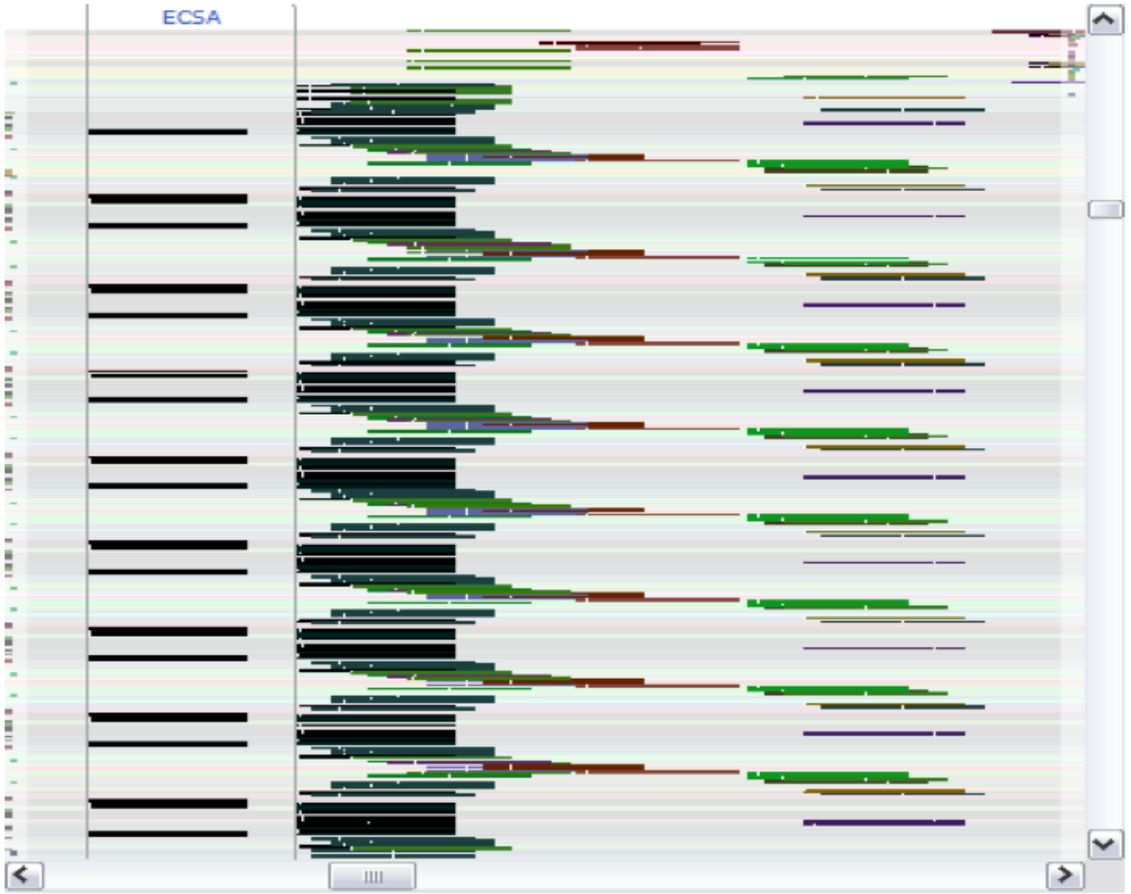
Roadmap

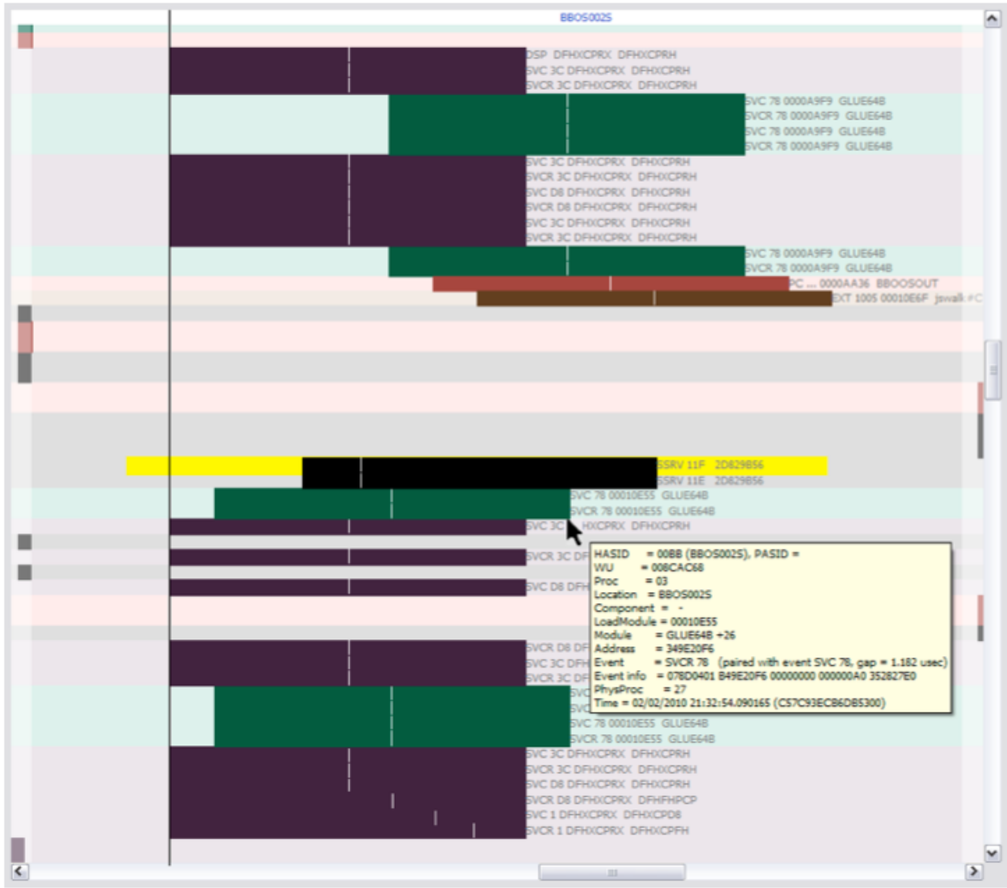
- Visual Perception
- Information Visualization
- **Software Visualization**
  - Structure
  - Evolution
  - **Behavior**

Zinsight is a visualization of large event traces using a *pixel based* representation



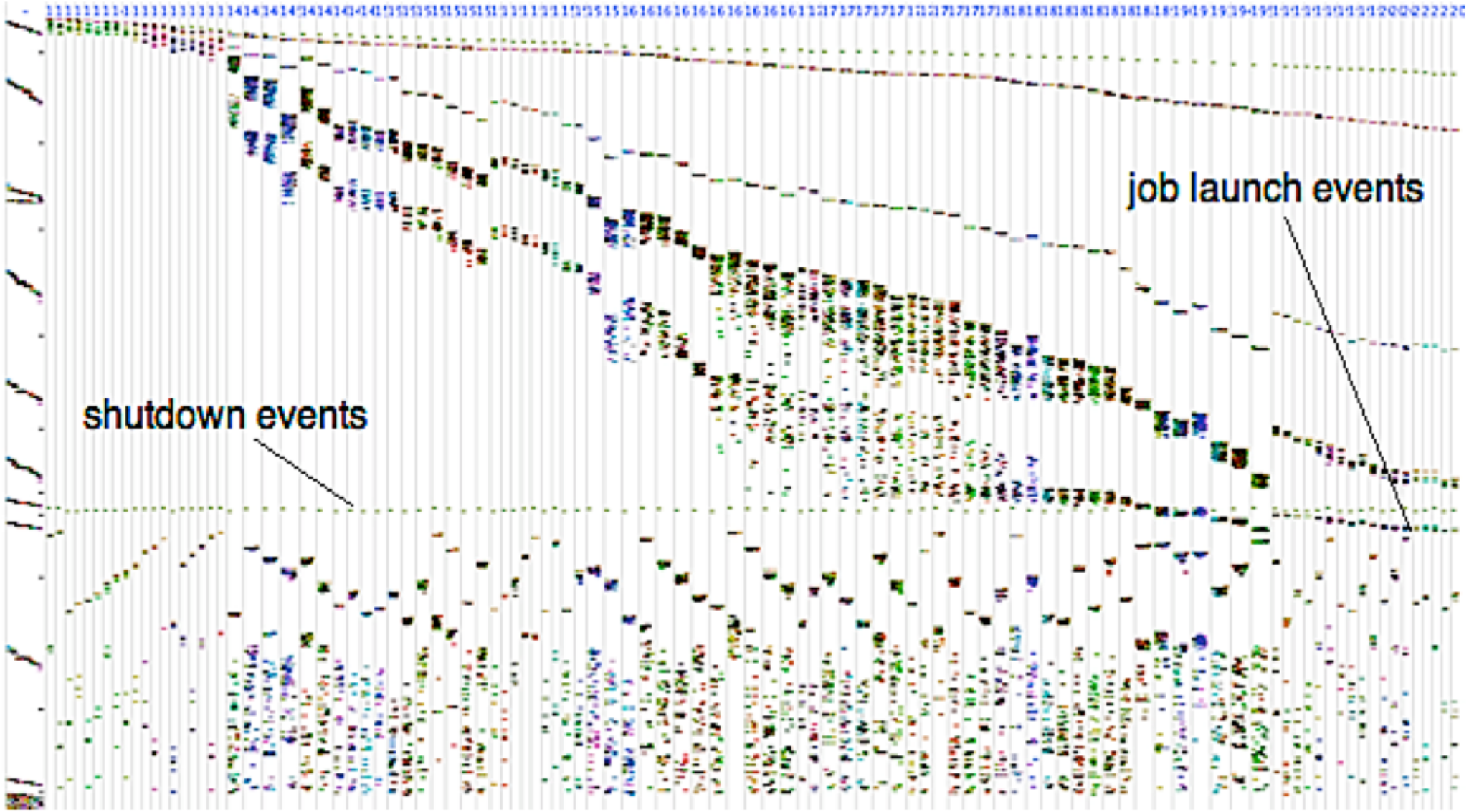
Massively reliant on visual pattern recognition and interactivity





Semantic Zooming

Visual detection of bugs

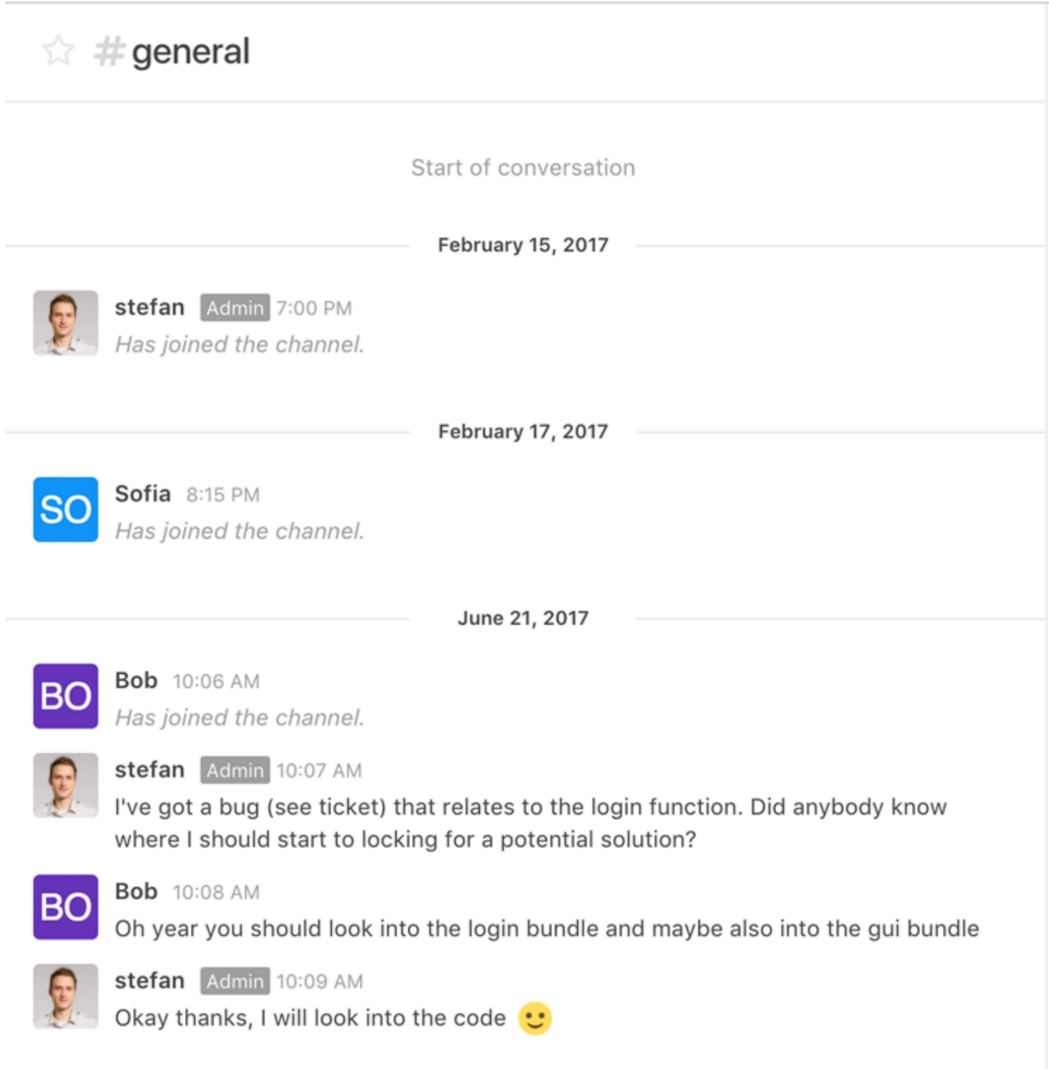


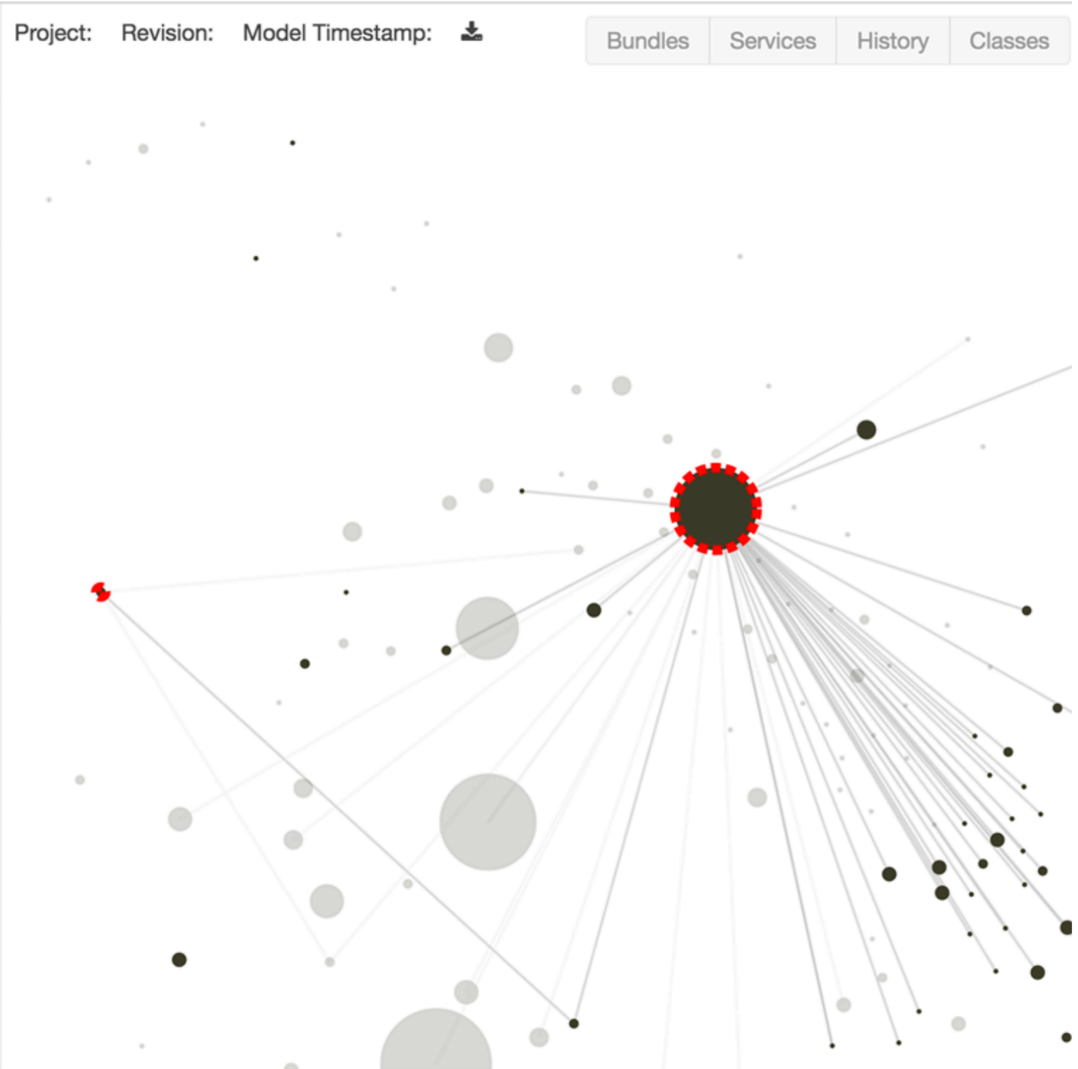
Present and Future of Software Visualization

- Novel interaction techniques

Present and Future of Software Visualization

Present and Future of Software Visualization





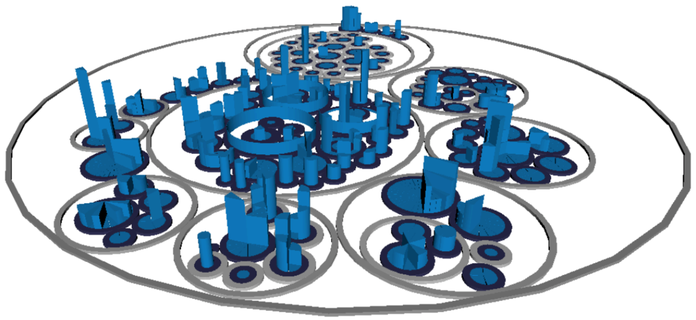
Present and Future of Software Visualization

- Novel interaction techniques
- **3D Software Visualization**

Present and Future of Software Visualization

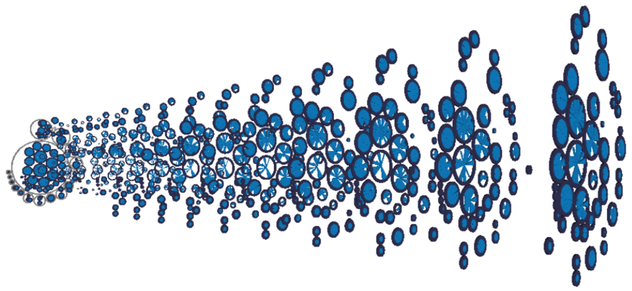
> Novel interaction techniques

> **3D Software Visualization**



Structure

Behavior



Evolution

Baum, D., Schilbach, J., Kovacs, P., Eisenecker, U., & Müller, R. (2017, September). GETAVIZ: generating structural, behavioral, and evolutionary views of software systems for empirical evaluation. In *Software Visualization (VISSOFT), 2017 IEEE Working Conference on* (pp. 114-118). IEEE.

Present and Future of Software Visualization

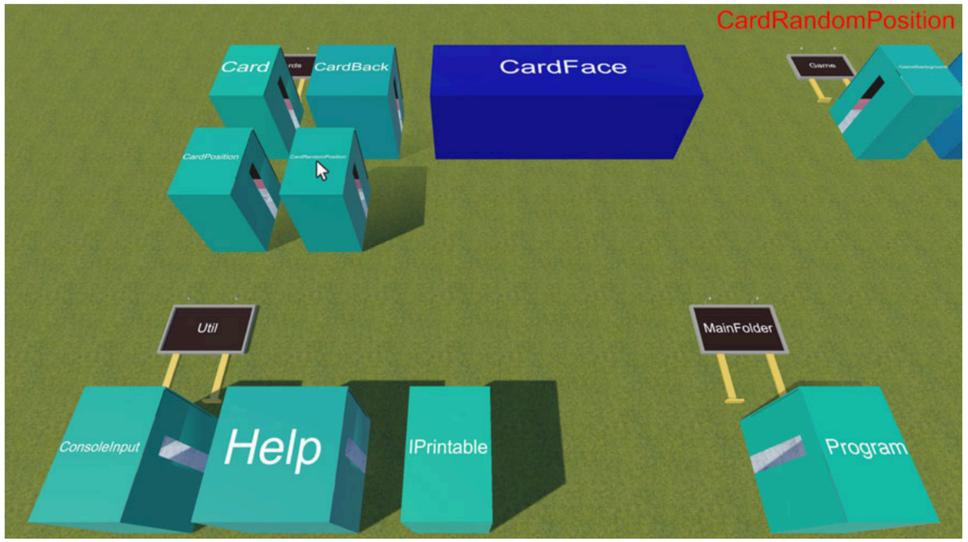
- Novel interaction techniques
- 3D Software Visualization
- **Multiple media to display visualizations**
  - **2D: Wall displays, Tabletops, Mobile devices**

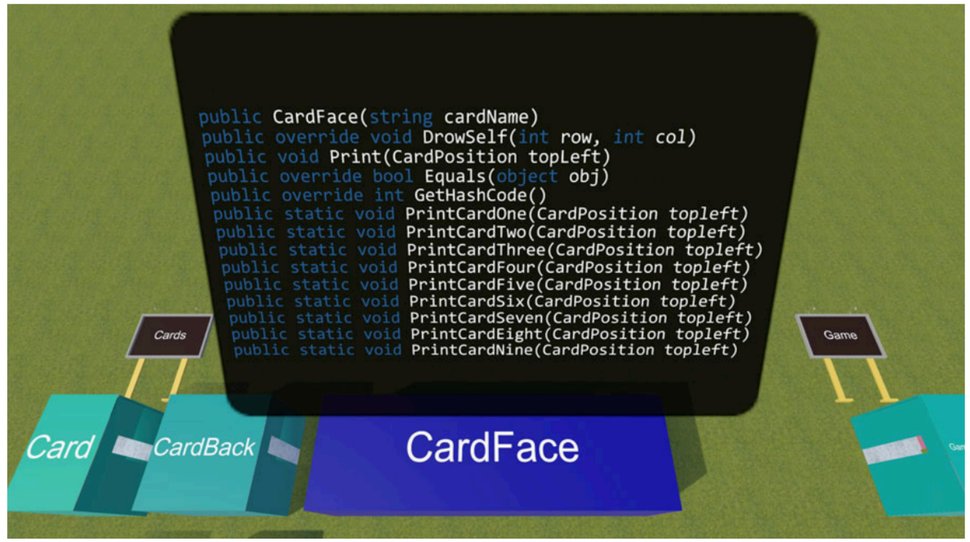

testling.exe#main
testling.exe#random_costs_test
testling.exe#printf
ucrtbase.dll#_stdio_common_vfprintf

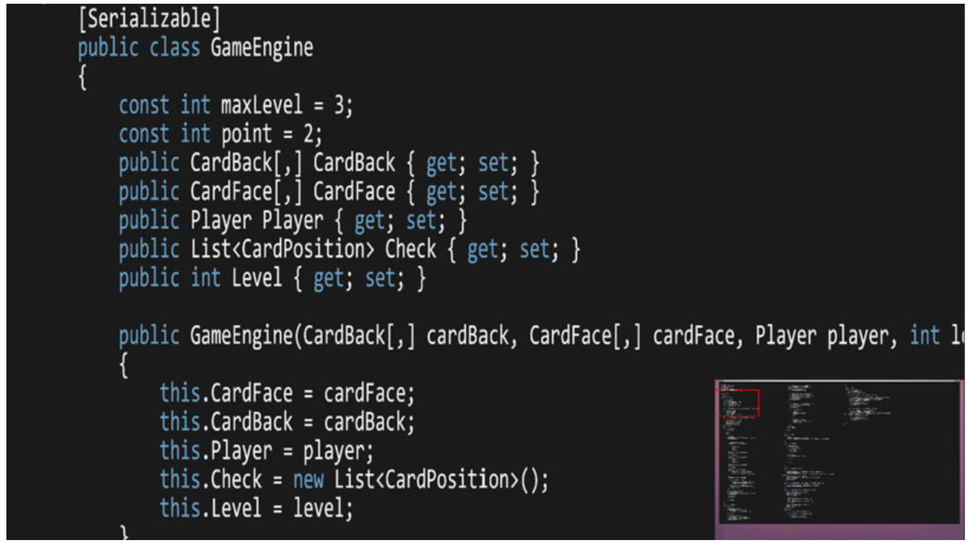
Present and Future of Software Visualization

- Novel interaction techniques
- 3D Software Visualization
- **Multiple media to display visualizations**
  - 2D: Wall displays, Tabletops, Mobile devices
  - **3D: AR, VR, CAVE, Stereo glasses, 3D prints**

Present and Future of Software Visualization

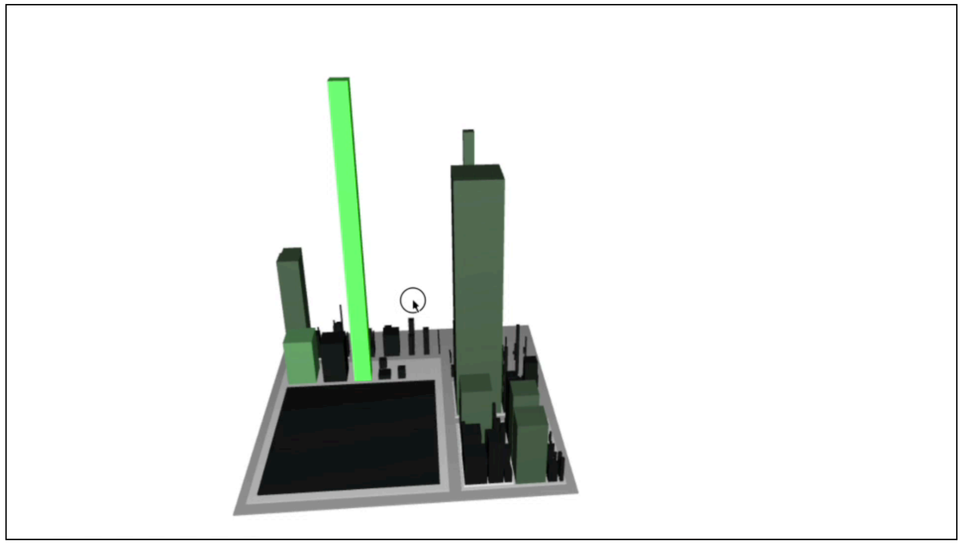


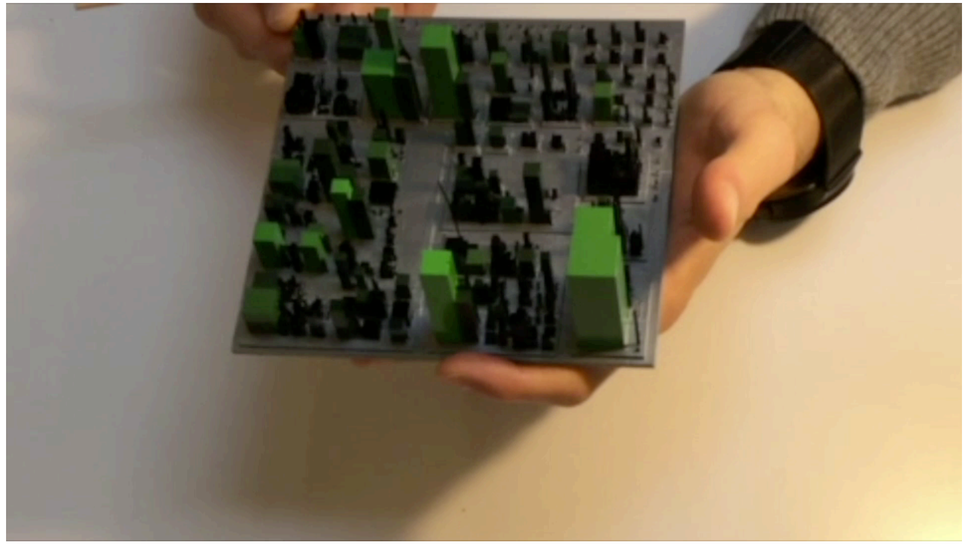


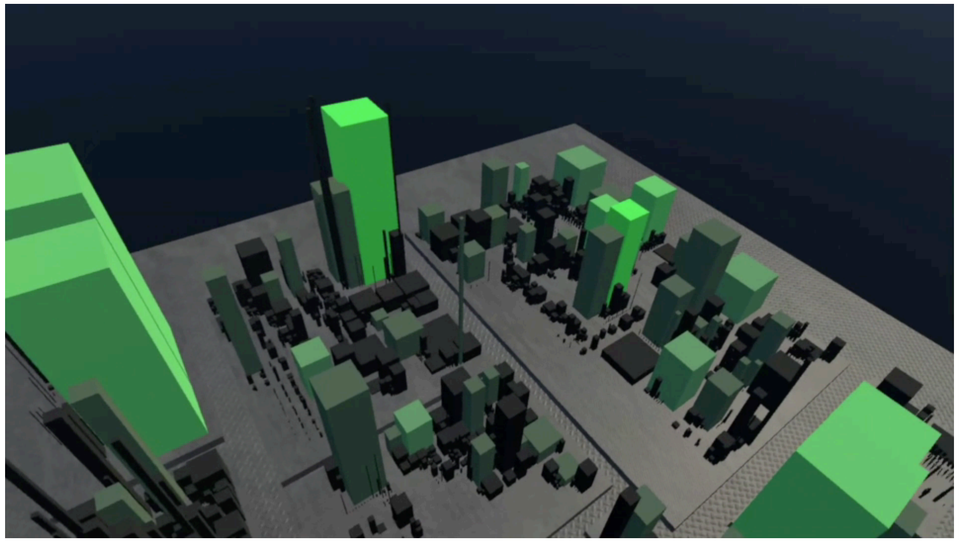


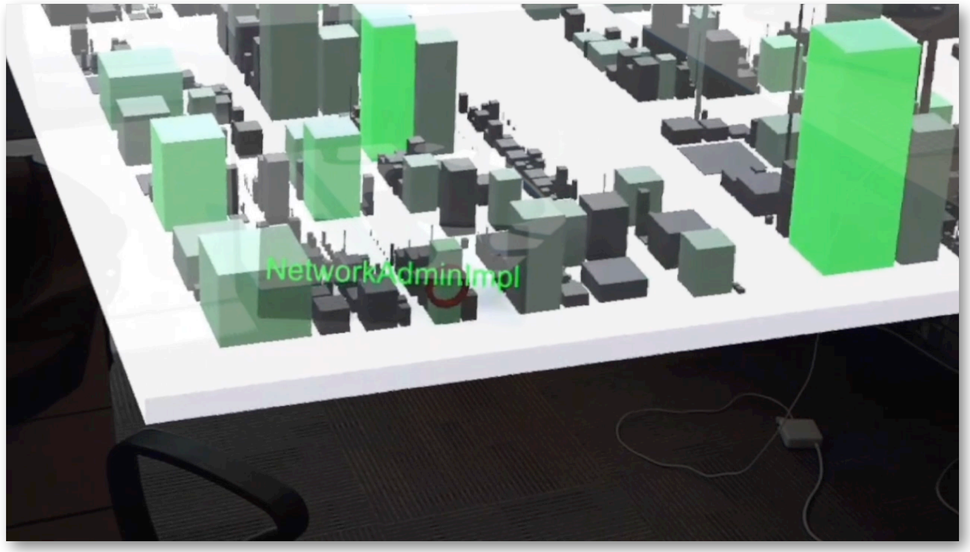
Khaloo, P., Maghoumi, M., Taranta, E., Bettner, D., & Laviola, J. (2017, September). Code Park: A New 3D Code Visualization Tool. In *Software Visualization (VISSOFT), 2017 IEEE Working Conference on* (pp. 43-53). IEEE.

Present and Future of Software Visualization



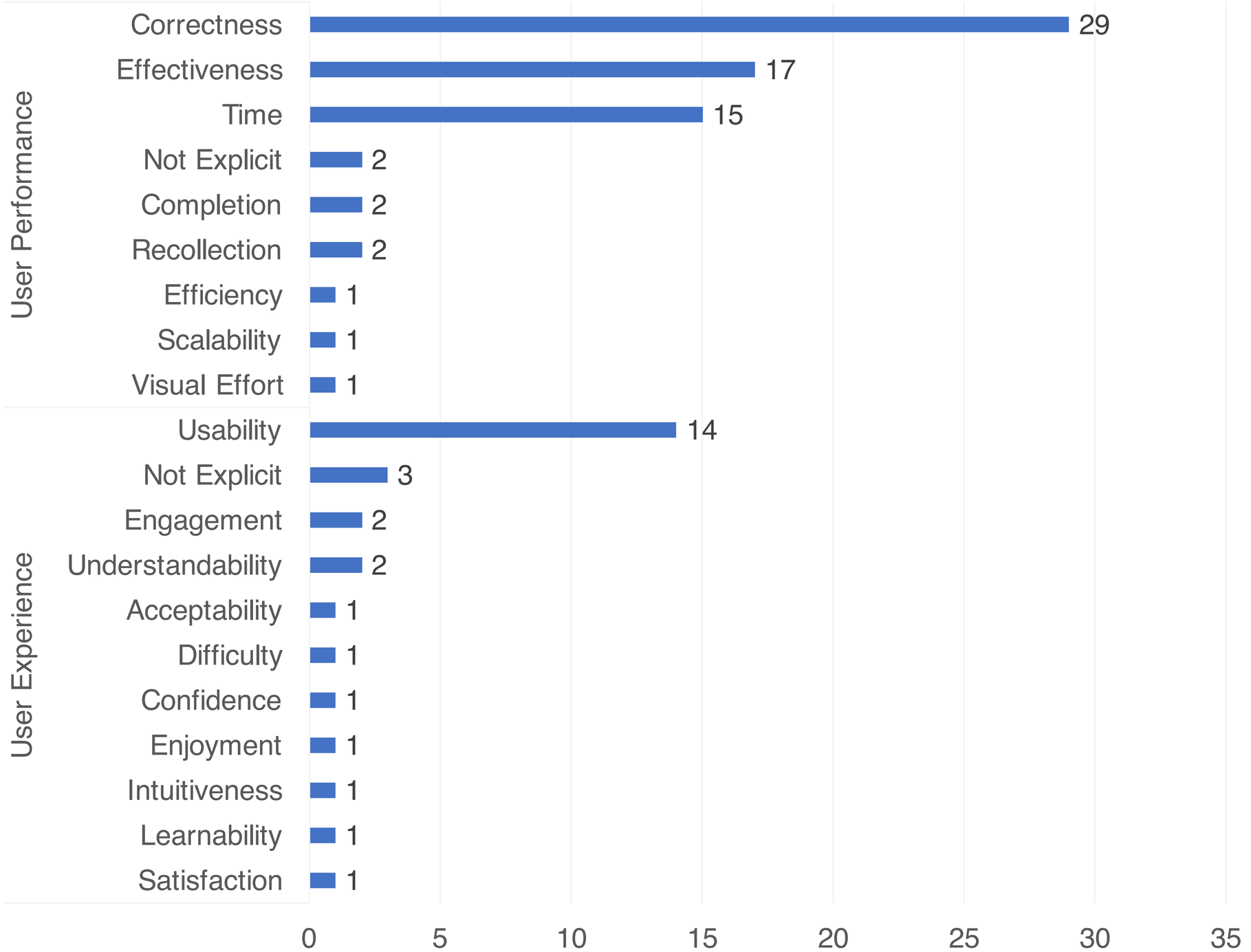






Present and Future of Software Visualization

- Novel interaction techniques
- 3D Software Visualization
- Multiple media to display visualizations
  - 2D: Wall displays, Tabletops, Mobile devices
  - 3D: AR, VR, CAVE, Stereo glasses, 3D prints
- **Focus on evaluation of software visualizations**
  - **Beyond time & Correctness**


User Performance
Correctness 29
Effectiveness 17
Time 15
Not Explicit 2
Completion 2
Recollection 2
Efficiency 1
Scalability 1
Visual Effort 1
User Experience
Usability 14
Not Explicit 3
Engagement 2
Understandability 2
Acceptability 1
Difficulty 1
Confidence 1
Enjoyment 1
Intuitiveness 1
Learnability 1
Satisfaction 1
0 5 10 15 20 25 30 35

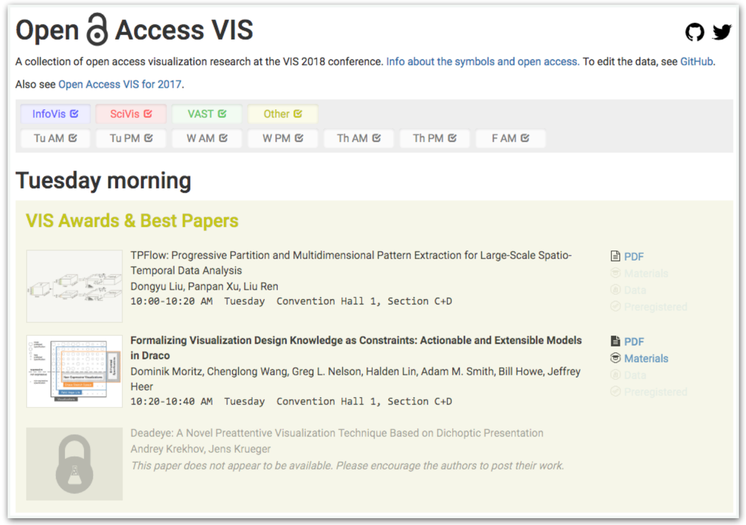
Present and Future of Software Visualization

- Novel interaction techniques
- 3D Software Visualization
- Multiple media to display visualizations
  - 2D: Wall displays, Tabletops, Mobile devices
  - 3D: AR, VR, CAVE, Stereo glasses, 3D prints
- **Focus on evaluation of software visualizations**
  - Beyond time & Correctness
  - **Replicability & Reproducibility**

Present and Future of Software Visualization

http://evaluate.inf.usi.ch/artifacts
https://www.artifact-eval.org/

https://www.acm.org/publications/policies/artifact-review-badging



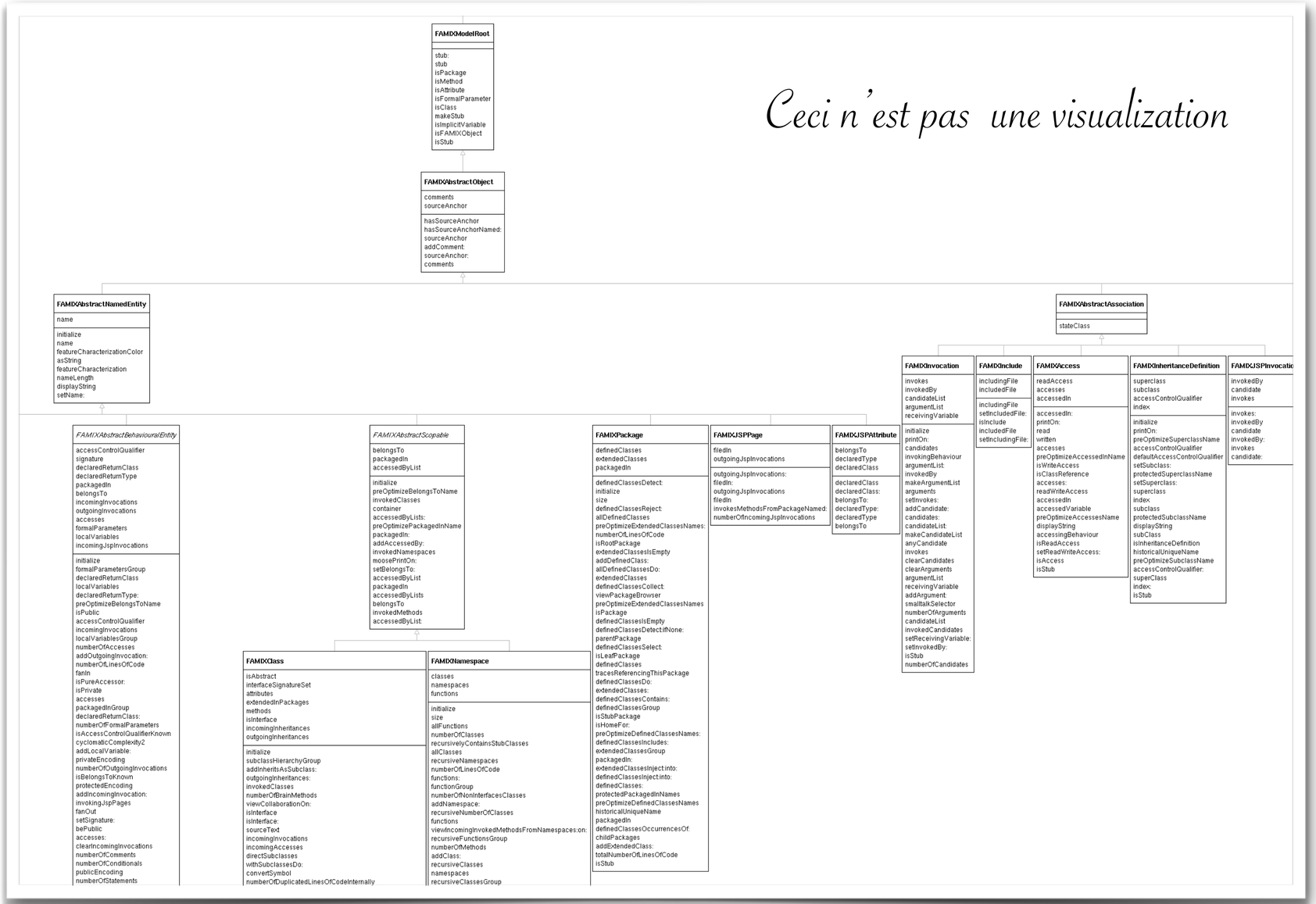
http://oavis.steveharoz.com/

Open Access Paper

Open Materials

Open Experiment Data

Inheritance



Ceci n'est pas une visualization

What you should know

- What is pre-attentive processing?
- What are the laws of Gestalt psychology?
- What is information visualization?
- Which aspects of software are usually visualized?
- Which techniques allow to visualize software structure?
- On what visualization features can we map evolution?
- What kinds of problems can be solved with software visualization?