b
UNIVERSITÄT
BERN

Software Visualization

Mircea Lungu

John Snow

The lines on the buildings are proportional to the number of deaths due to cholera from those buildings.

Do you know how did Snow realize which was the cause of cholera by looking at this map?

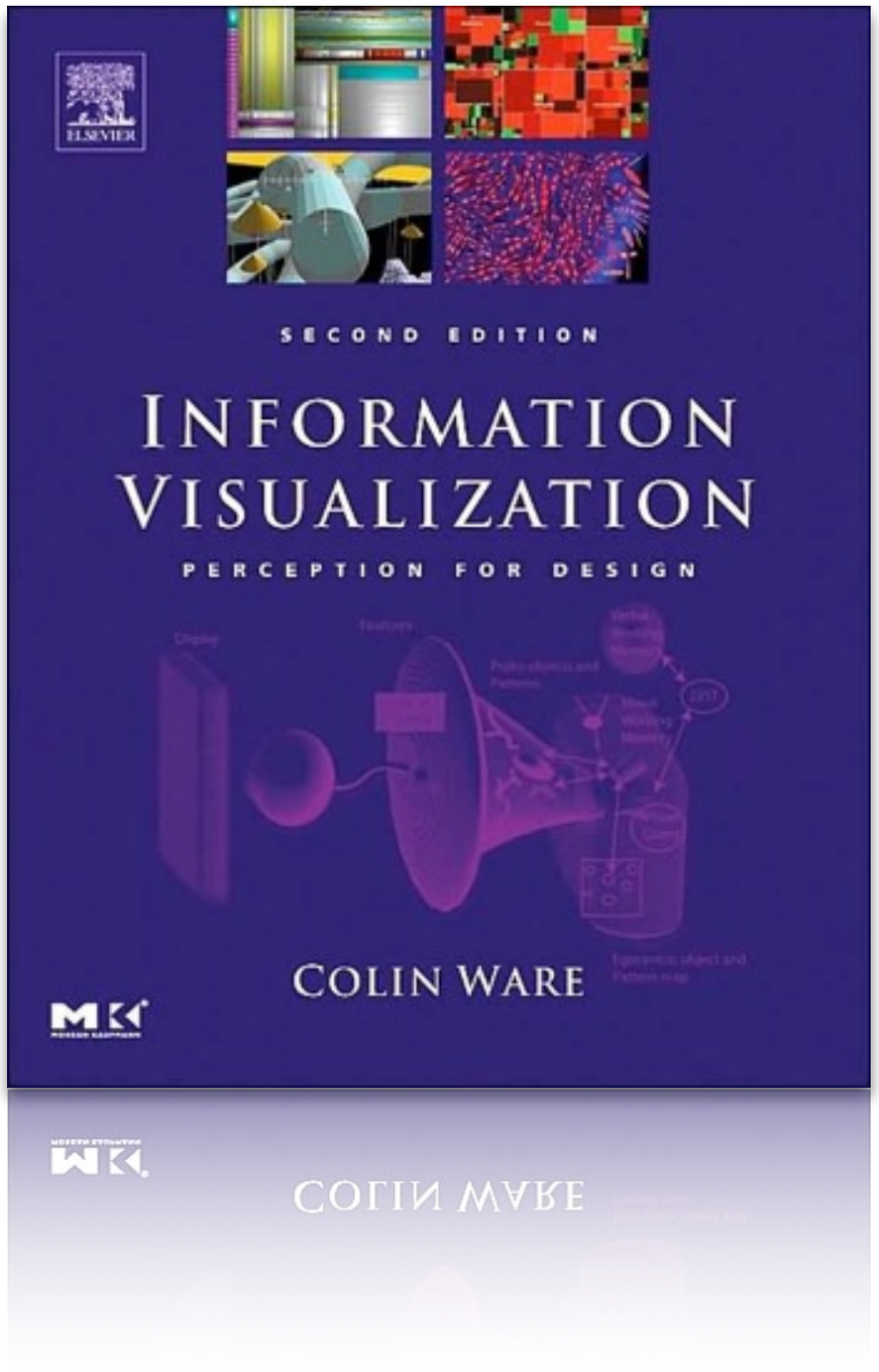
Roadmap

- Visual Perception
- Information Visualization
- Software Visualization

Roadmap

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

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



Preattentive Processing

Orientation	Line Length	Line Width	Size
Shape	Curvature	Added Marks	Enclosure

Preattentive Processing: Color

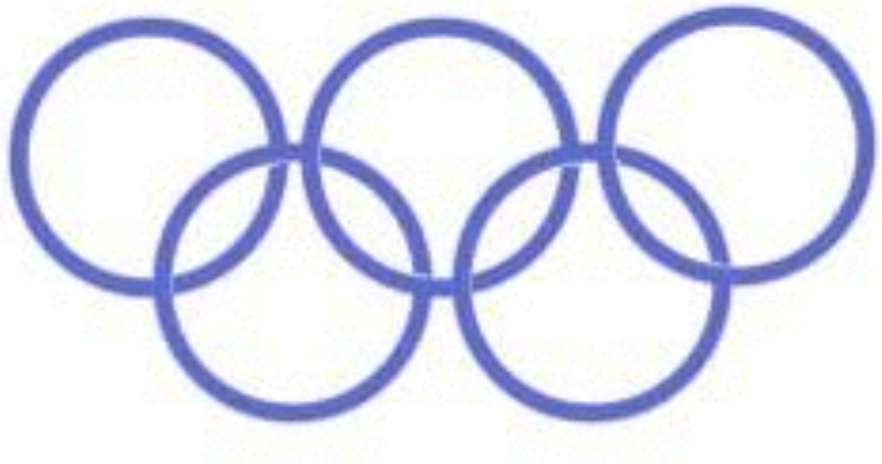
8789364082376403128764532984732984732094873290845389274-0329874-32874-2319847509834098340983240983204982-0984903281453209481-0839393947896587436598

Preattentive Processing: Color

8789364082376403128764532984732984732094873290845 3
89274-0329874-32874-2319847509834098340983240983204
9823-0984903281453209481-0839393947896587436598

Gestalt Psychology

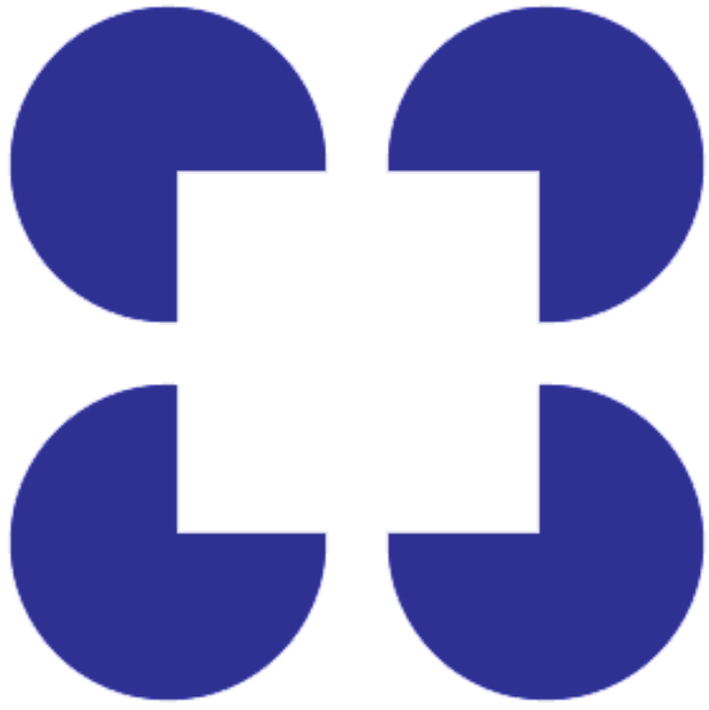
- The law of simplicity
- The Gestalt Laws
  1. Closure
  2. Similarity
  3. Proximity
  4. Continuity



“Reality is organized and reduced to the simplest form possible”

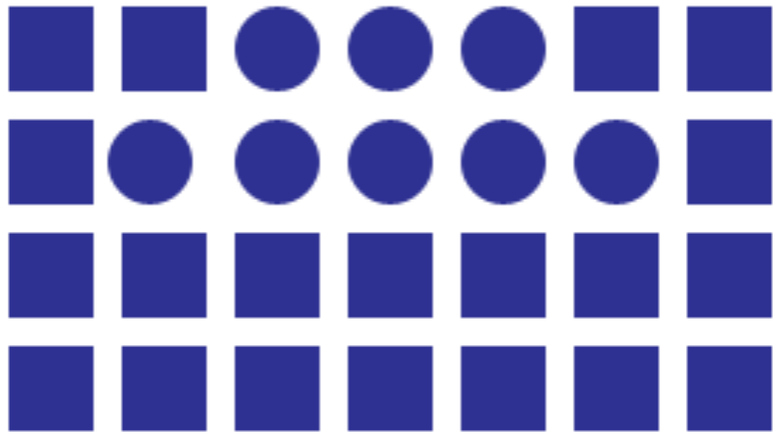
1. Law of Closure

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

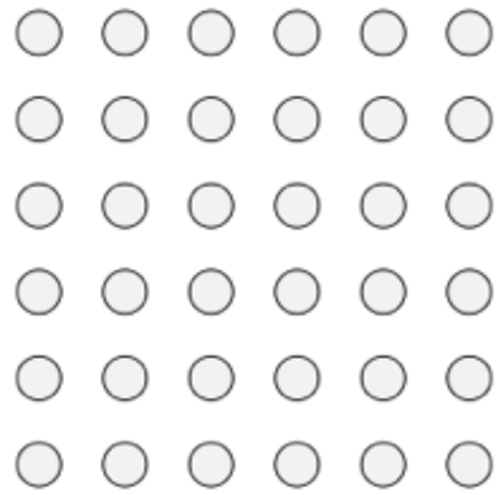


2. Law of Similarity

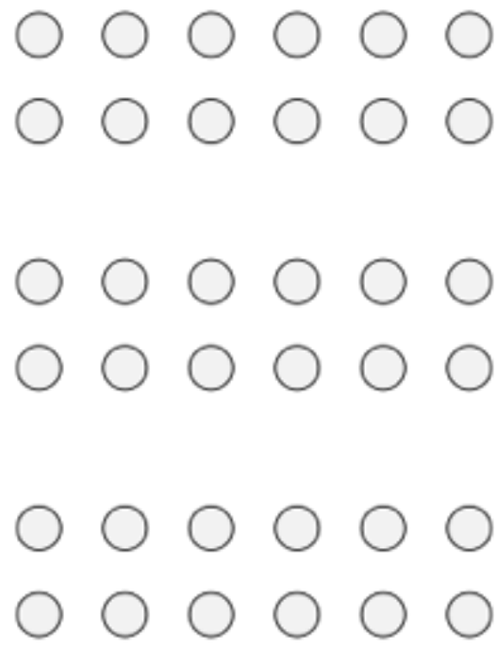
- The mind groups similar elements together



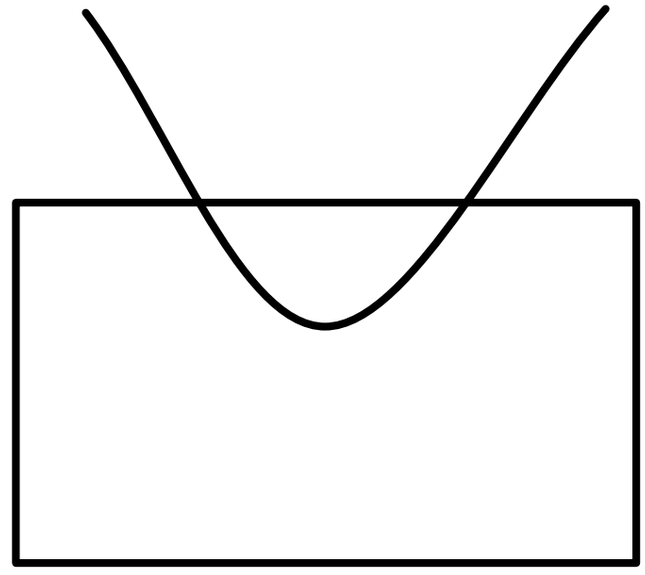
3. Law of Proximity



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



4. Law of Continuity



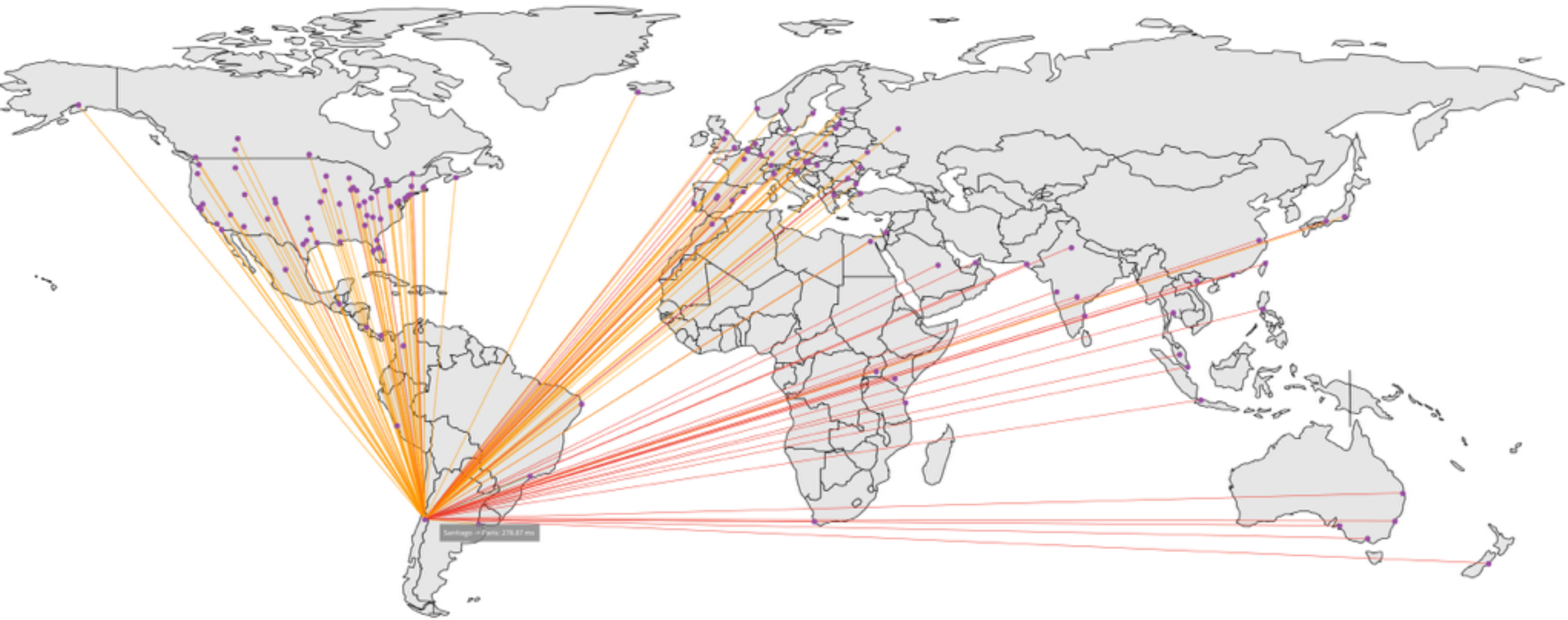
> Lines follow the smoothest and simplest path.

Roadmap

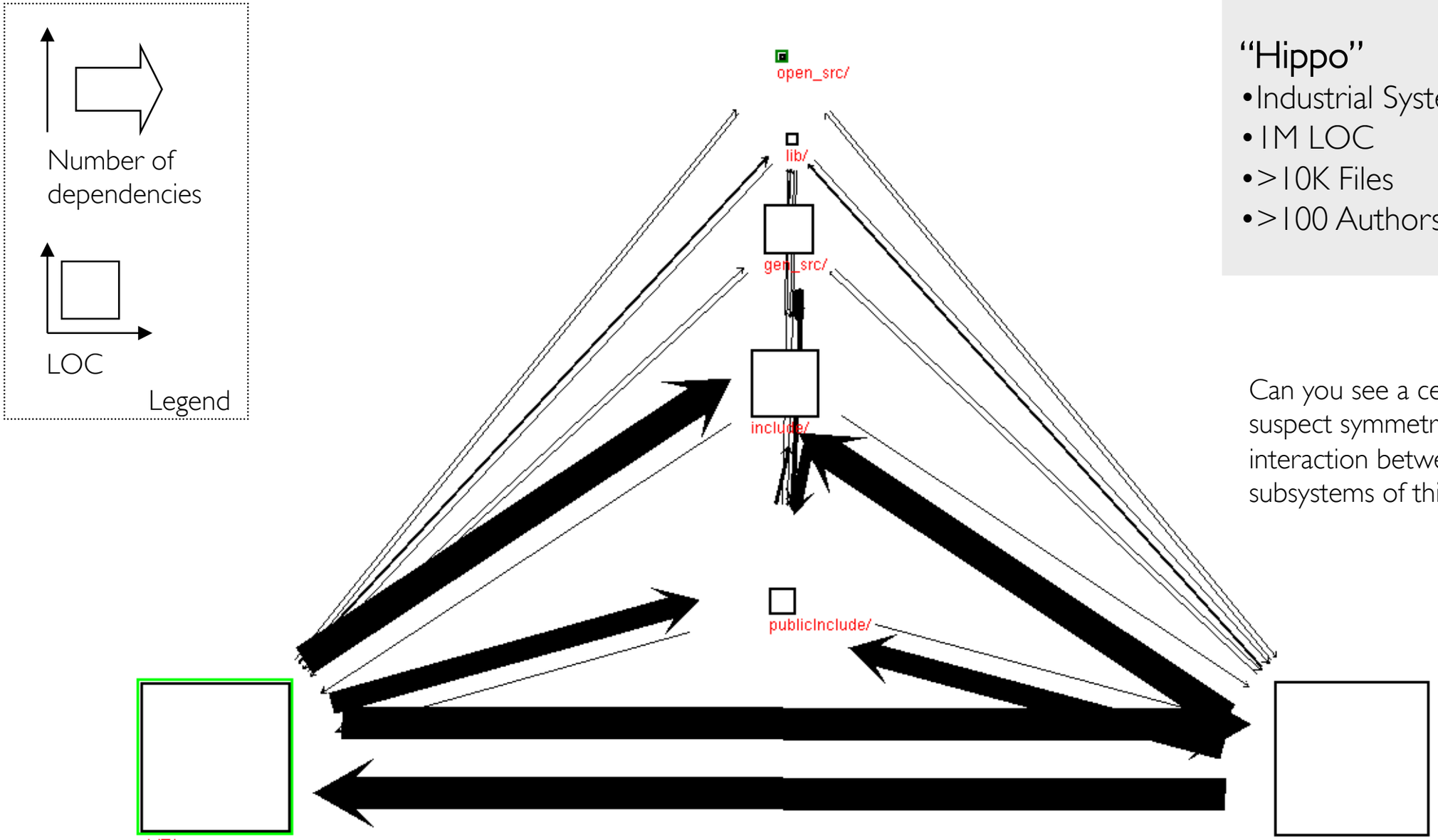
- Visual Perception
- **Information Visualization**
- Software Visualization

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

Uncovers emergent properties and outliers



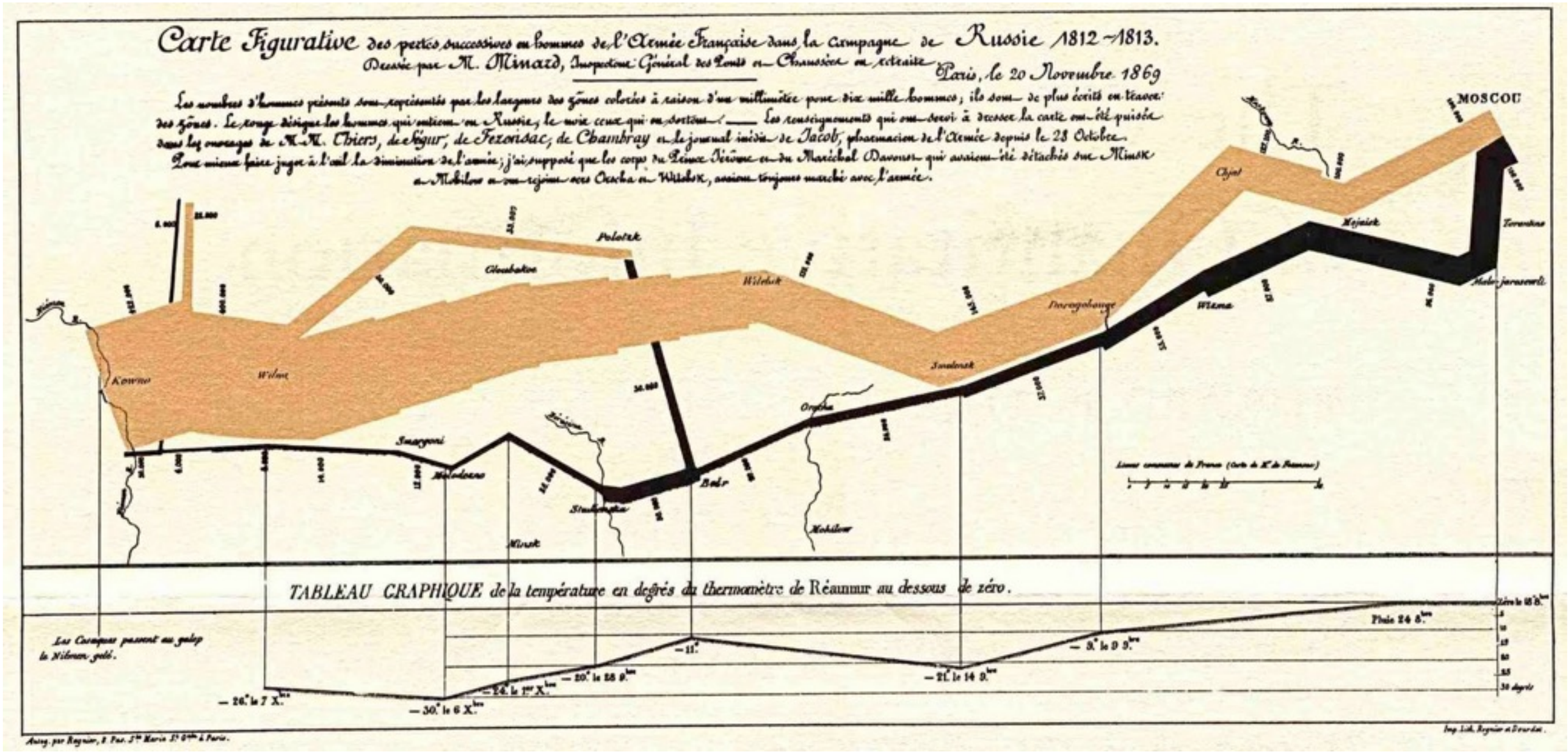
Exposes problems with the dataset



"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



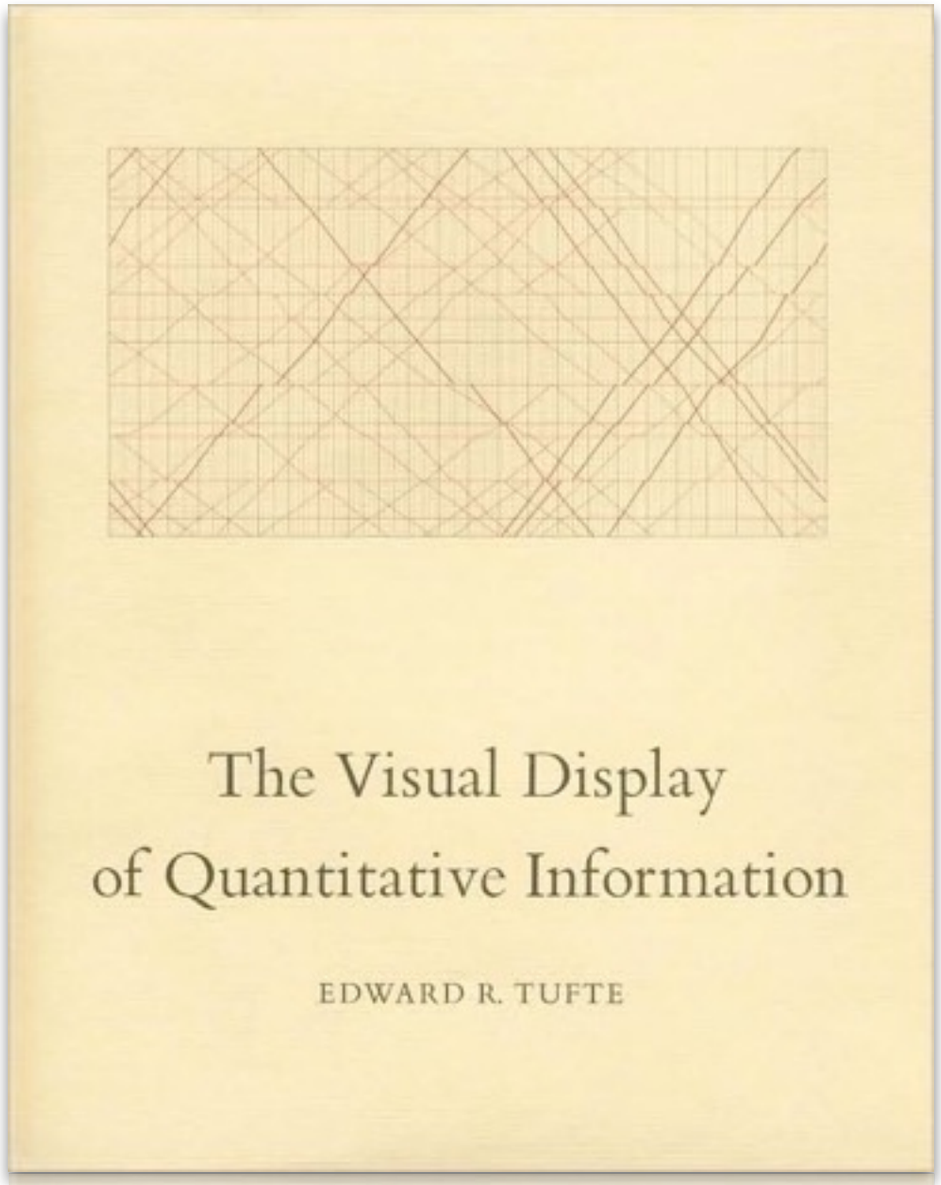
Uses of Information Visualization

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

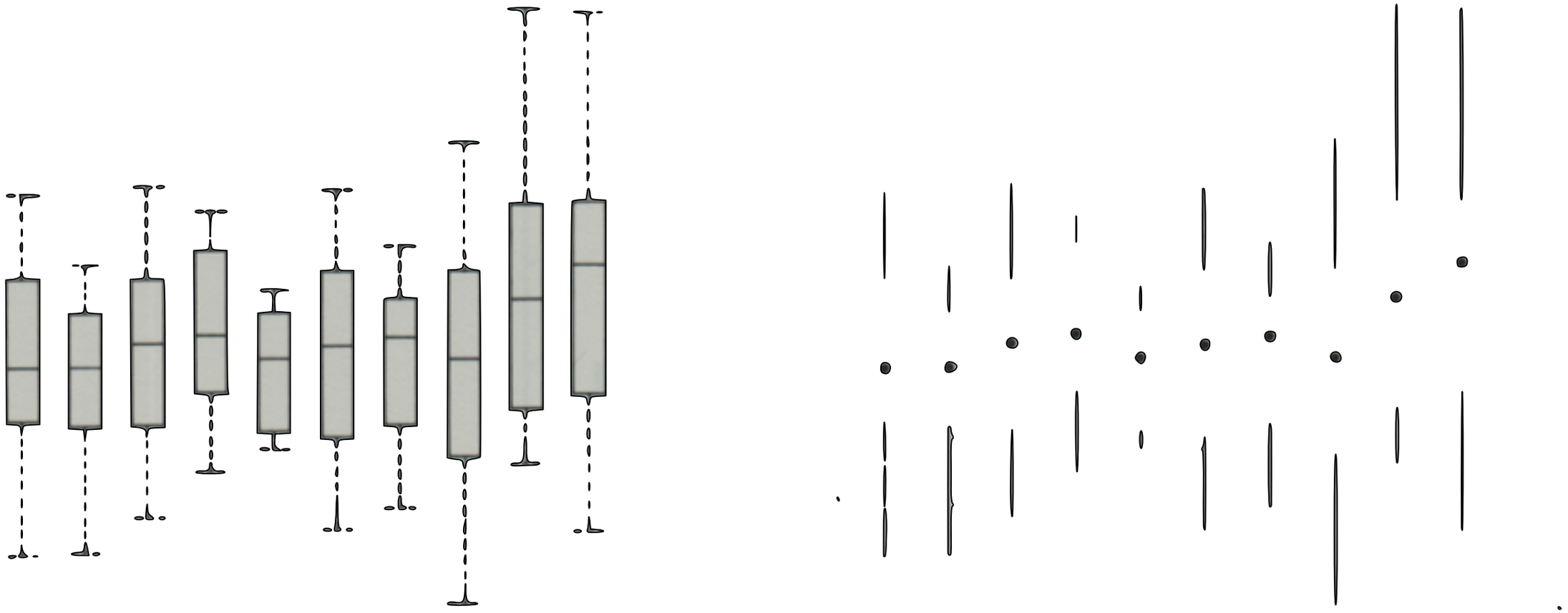
Roadmap

- Visual Perception
- Information Visualization
  - **Visualization Principles [Intermezzo]**
- Visualizing Software

Good information visualization is based on **style**, **integrity,** and **design.**

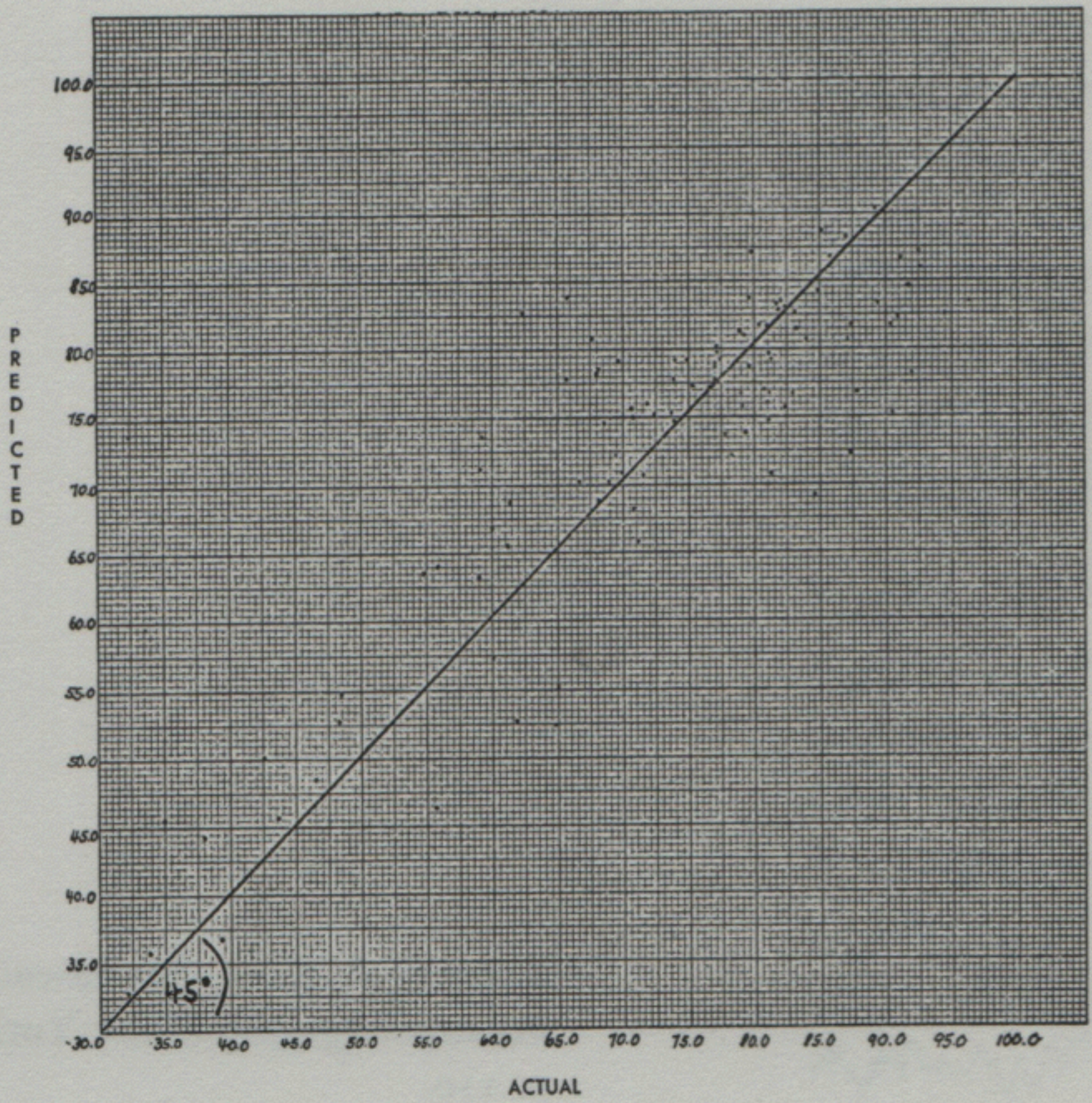


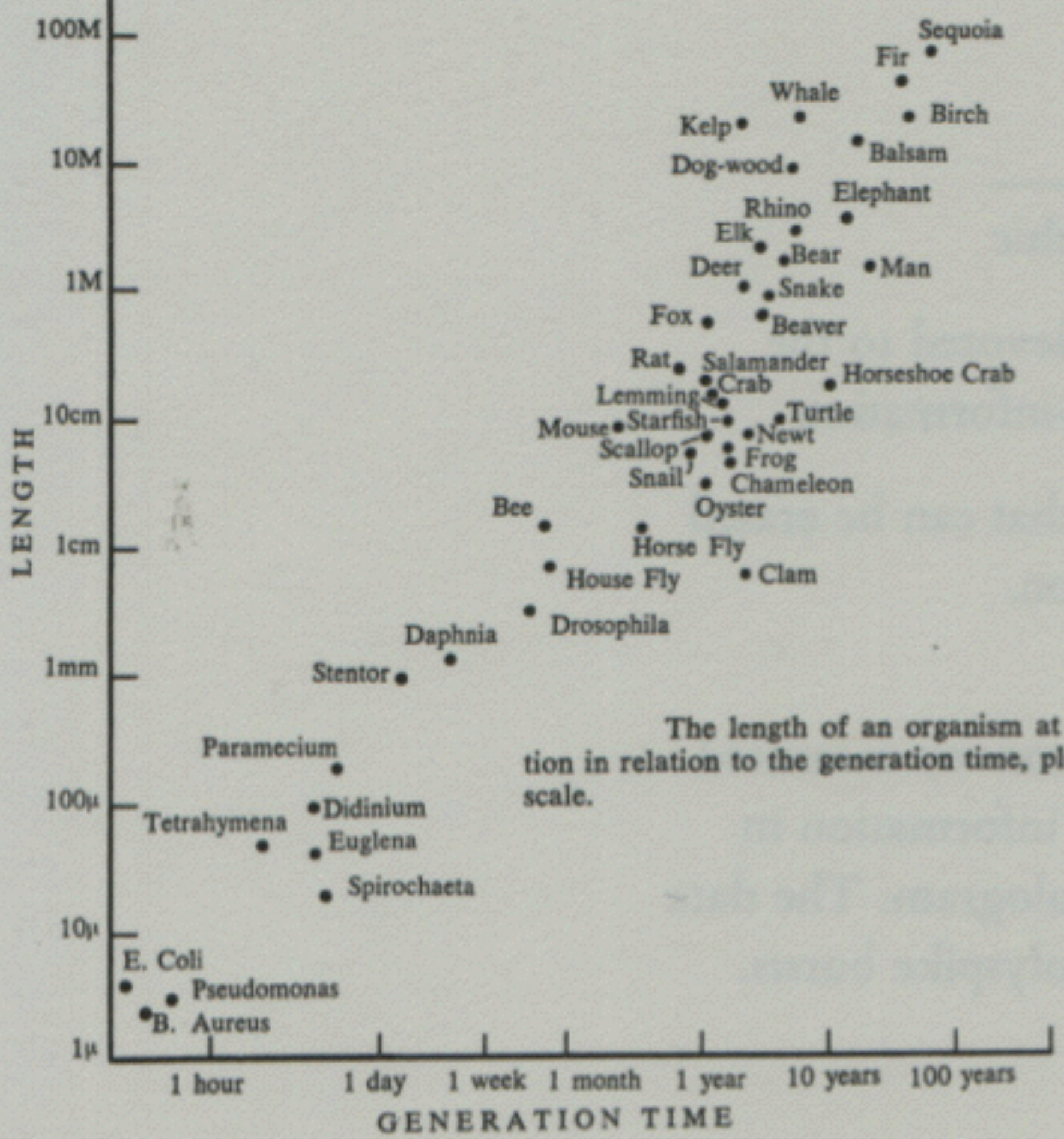
Style: Minimize Non-Data Ink



Removing ink from your graph should remove meaning from it.

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



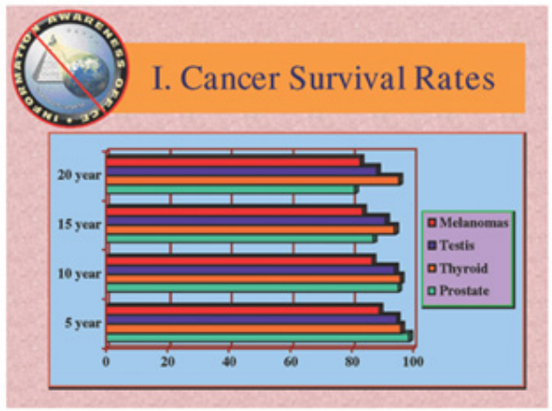


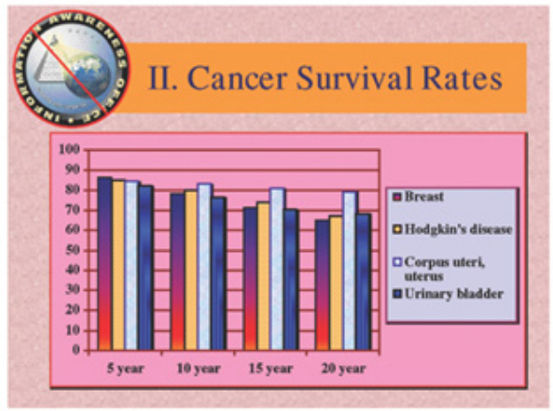
The length of an organism at the time of reproduction in relation to the generation time, plotted on a logarithmic scale.

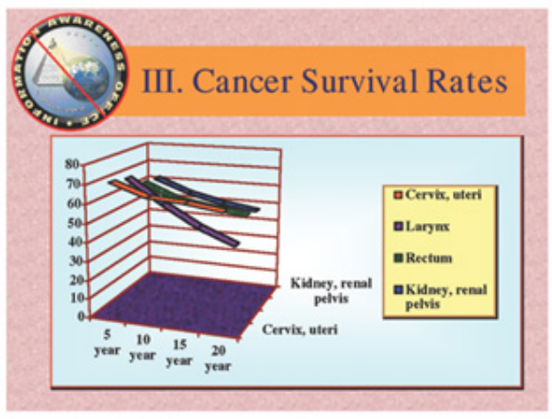
Design: Choose the appropriate representation

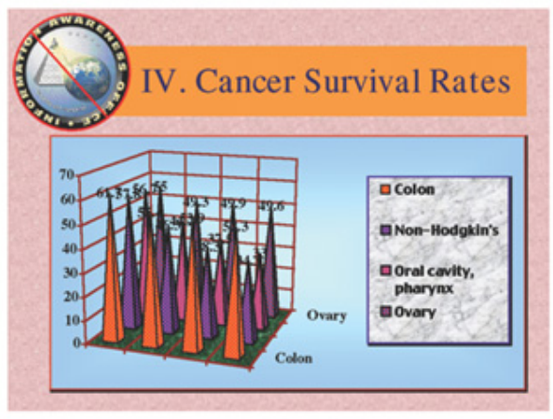
Estimates of relative survival rates, by cancer site

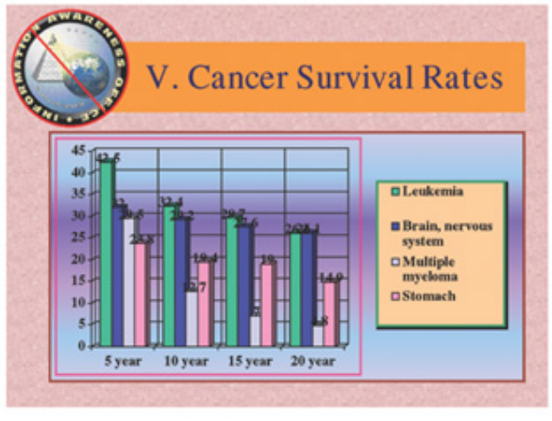
	% survival rates and standard errors							
	5 year		10 year		15 year		20 year	
Prostate	98.8	0.4	95.2	0.9	87.1	1.7	81.1	3.0
Thyroid	96.0	0.8	95.8	1.2	94.0	1.6	95.4	2.1
Testis	94.7	1.1	94.0	1.3	91.1	1.8	88.2	2.3
Melanomas	89.0	0.8	86.7	1.1	83.5	1.5	82.8	1.9
Breast	86.4	0.4	78.3	0.6	71.3	0.7	65.0	1.0
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

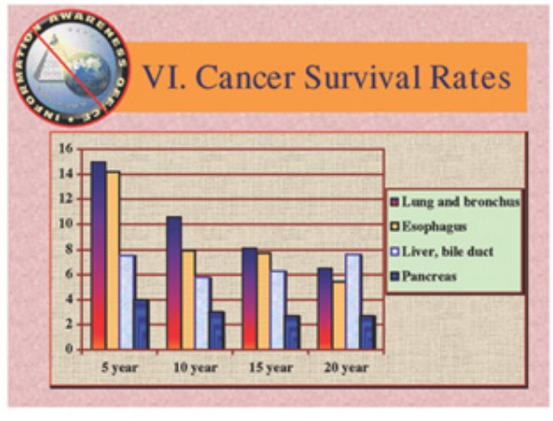




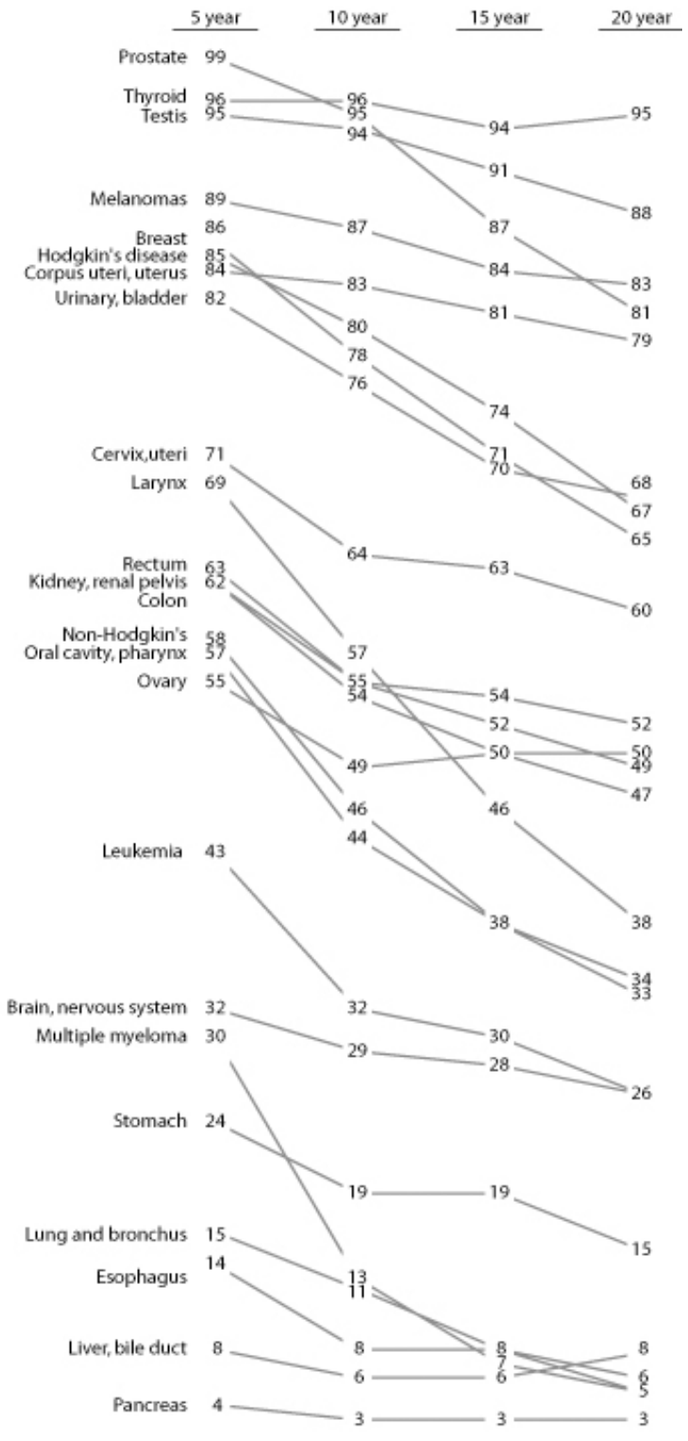




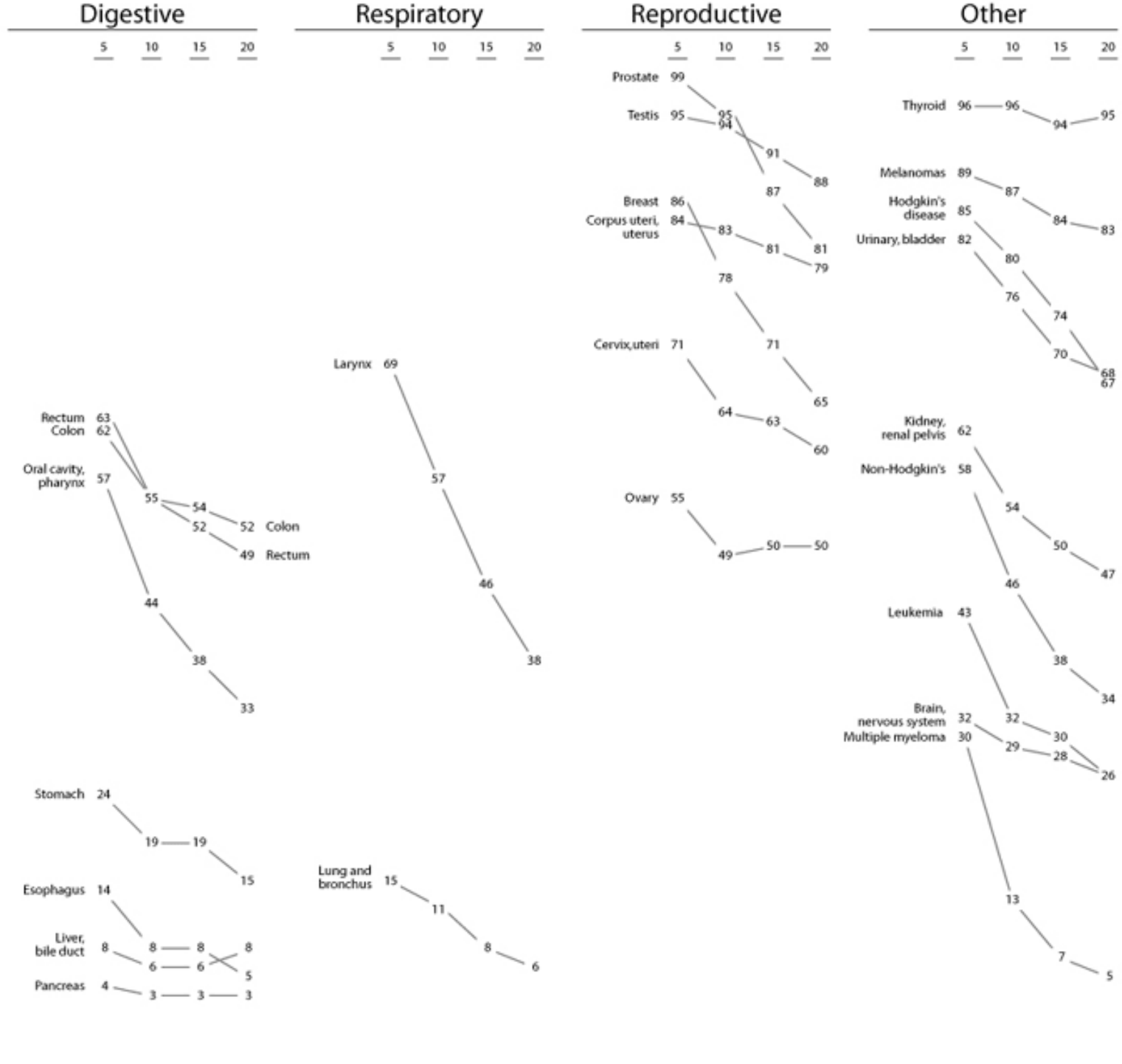




Estimates of relative survival rates, by cancer site

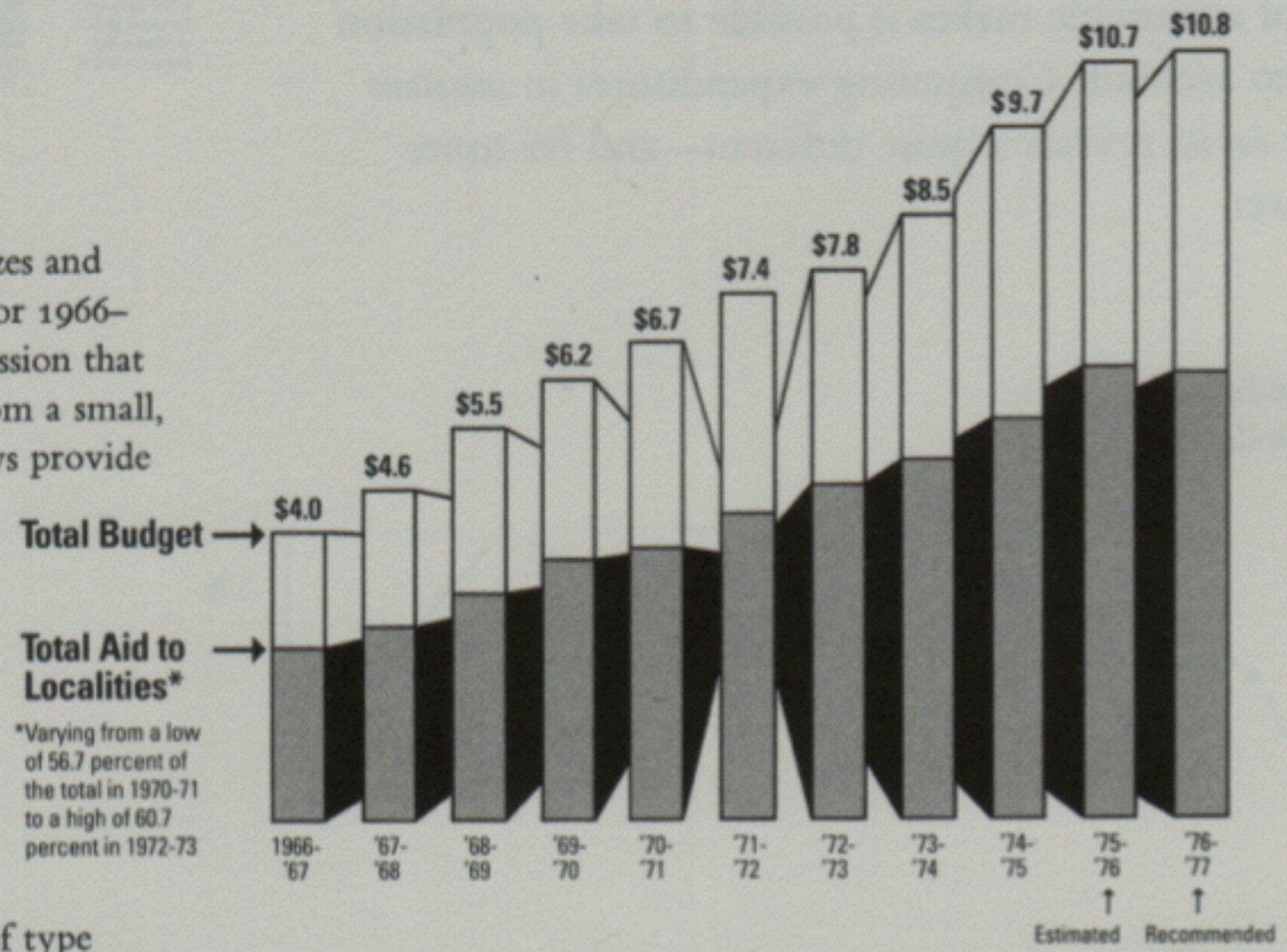


Estimates of relative survival rates, by cancer site



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.

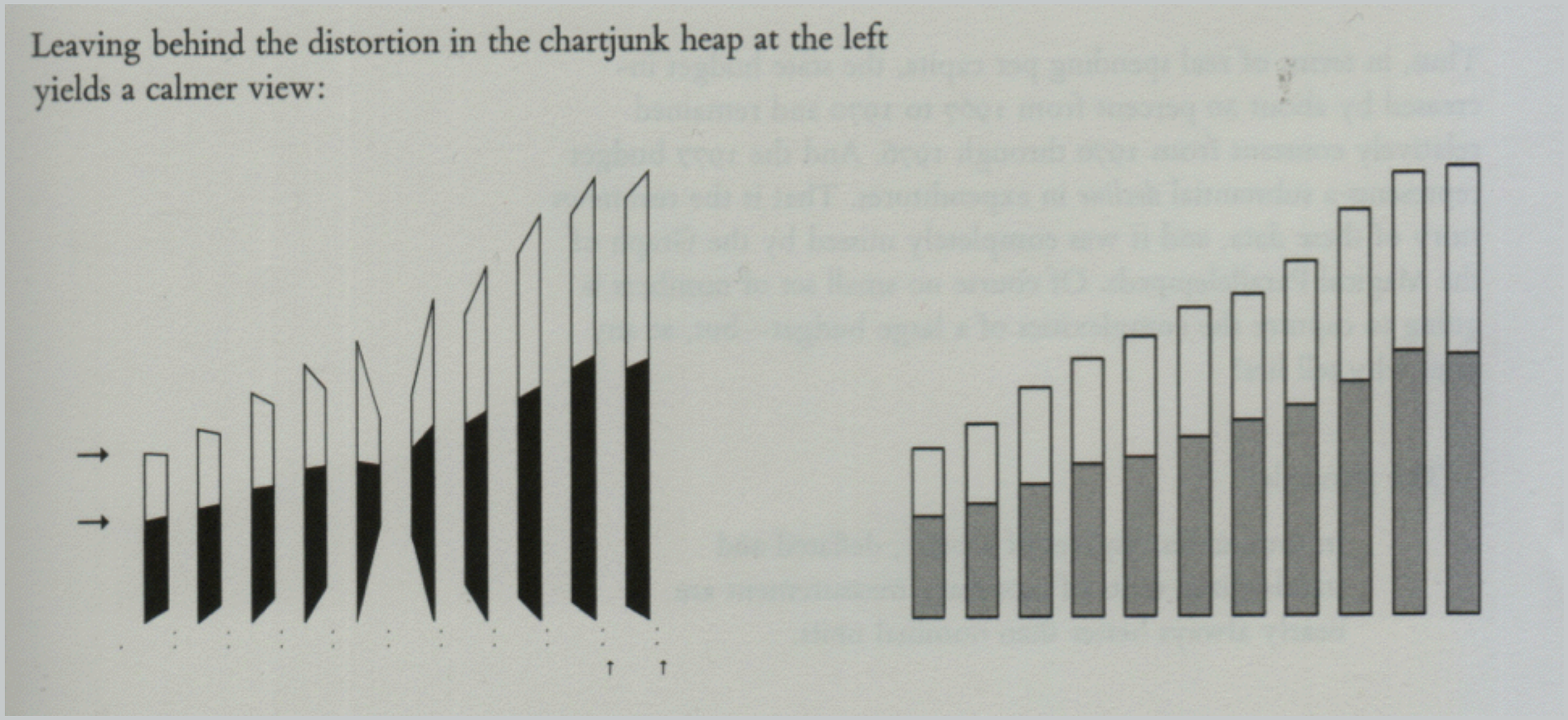


This squeezed-down block of type contributes to an image of small, squeezed-down budgets back in the good old days.

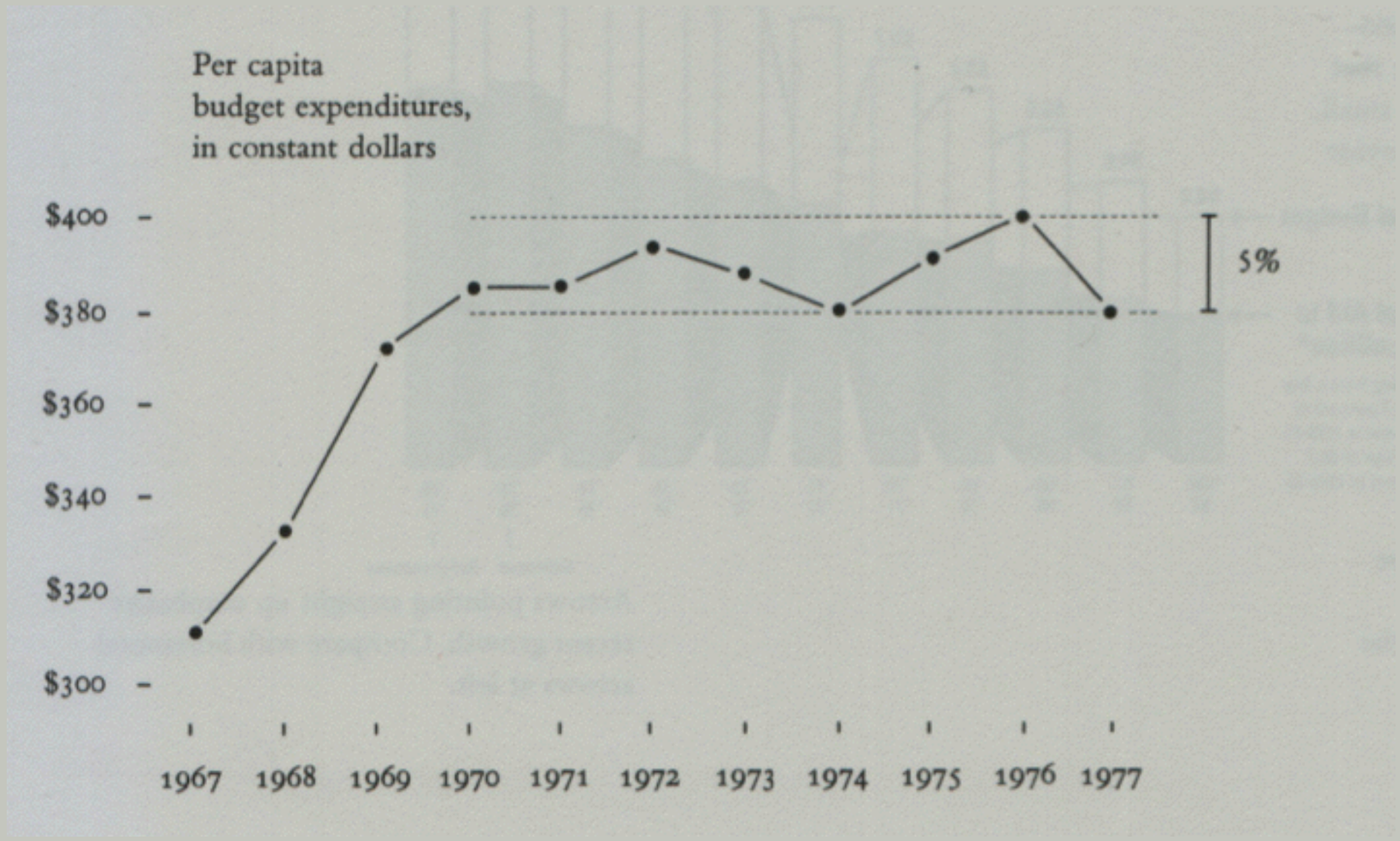
Arrows pointing straight up emphasize recent growth. Compare with horizontal arrows at left.

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

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

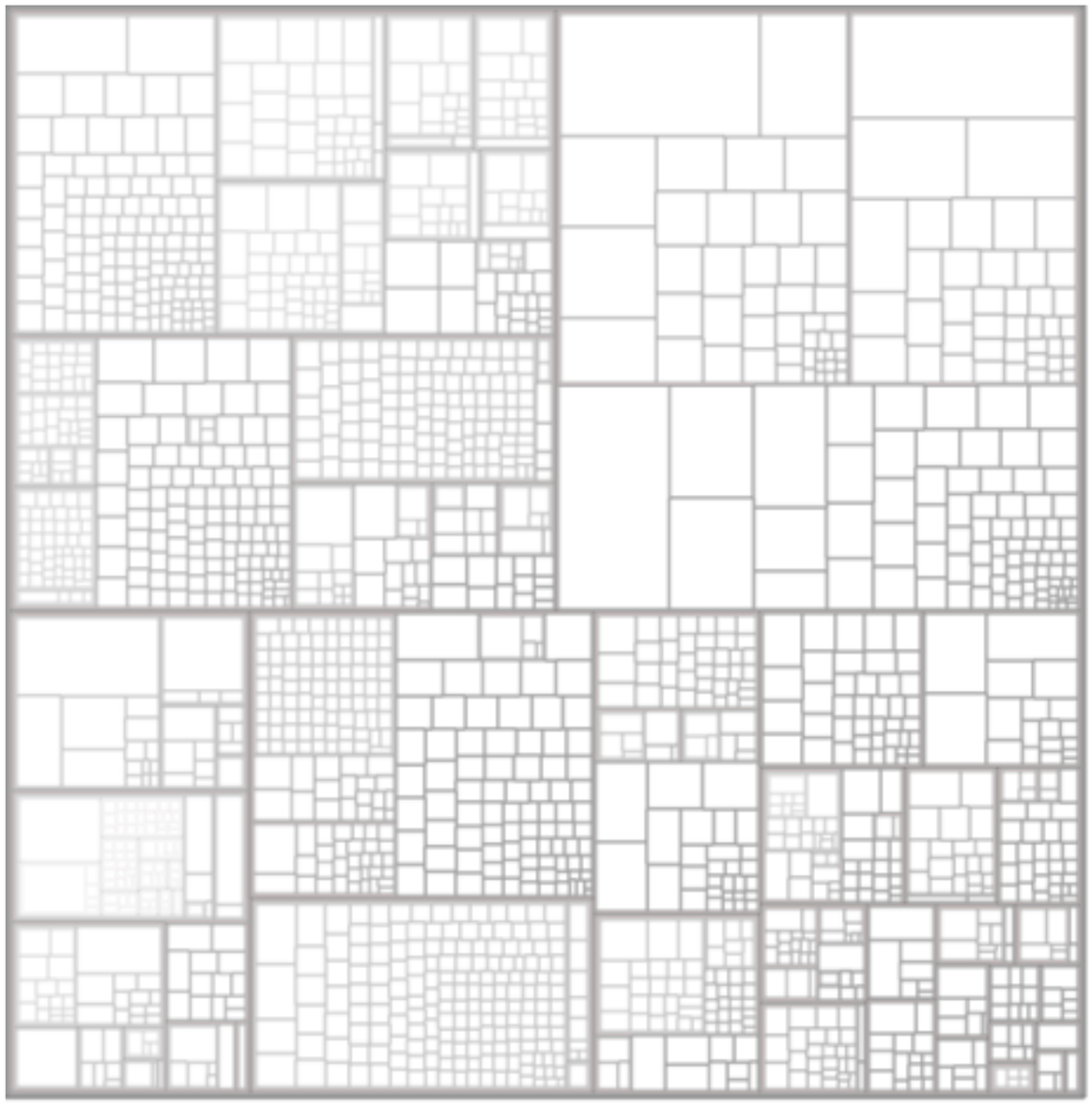
Roadmap

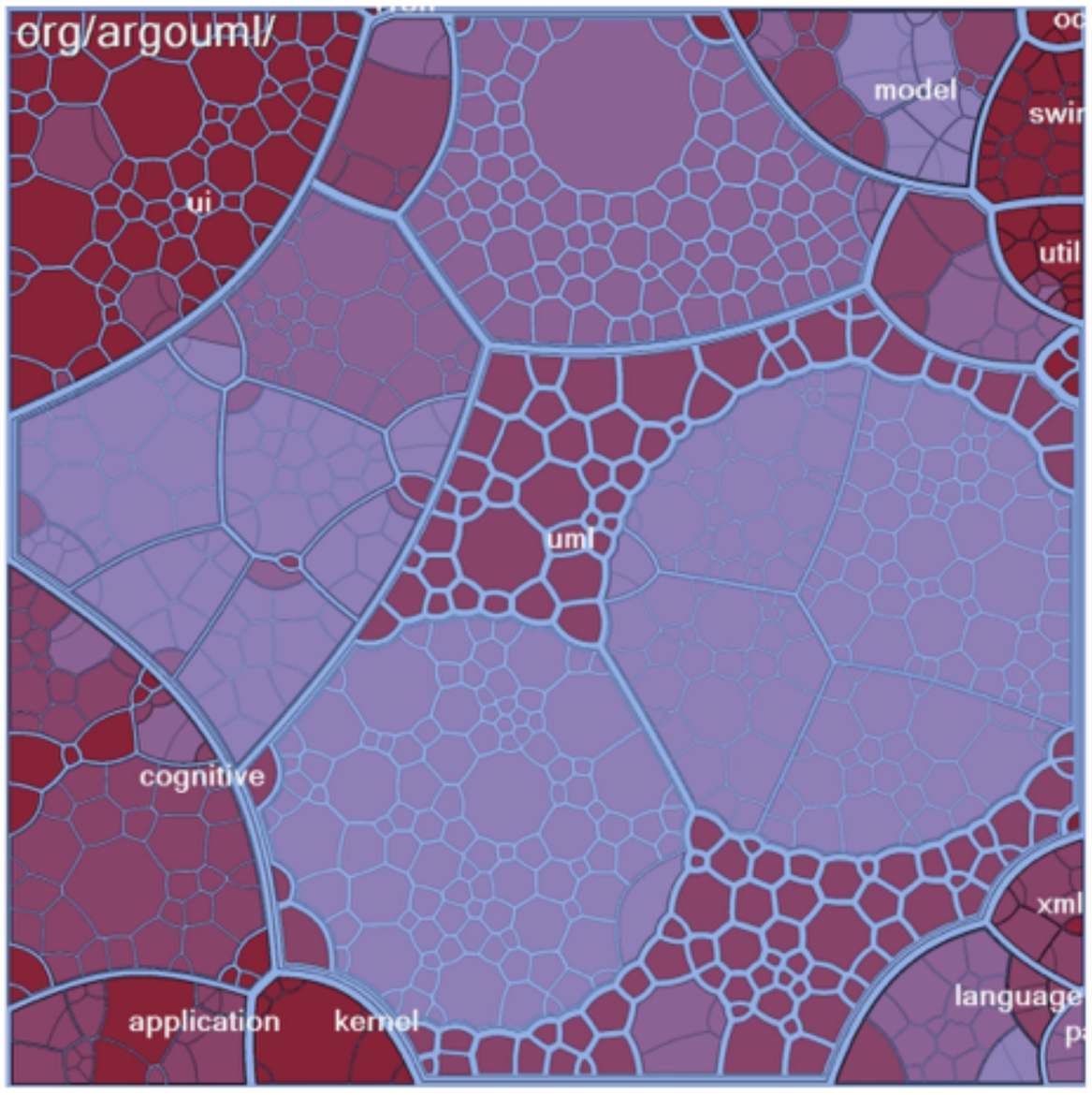
- Information Visualization
- Designing Visualizations
- **Software Visualization**
  - **Structure**
  - Evolution
  - Behavior

Space Filling Techniques

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

Providing an overview of size distribution

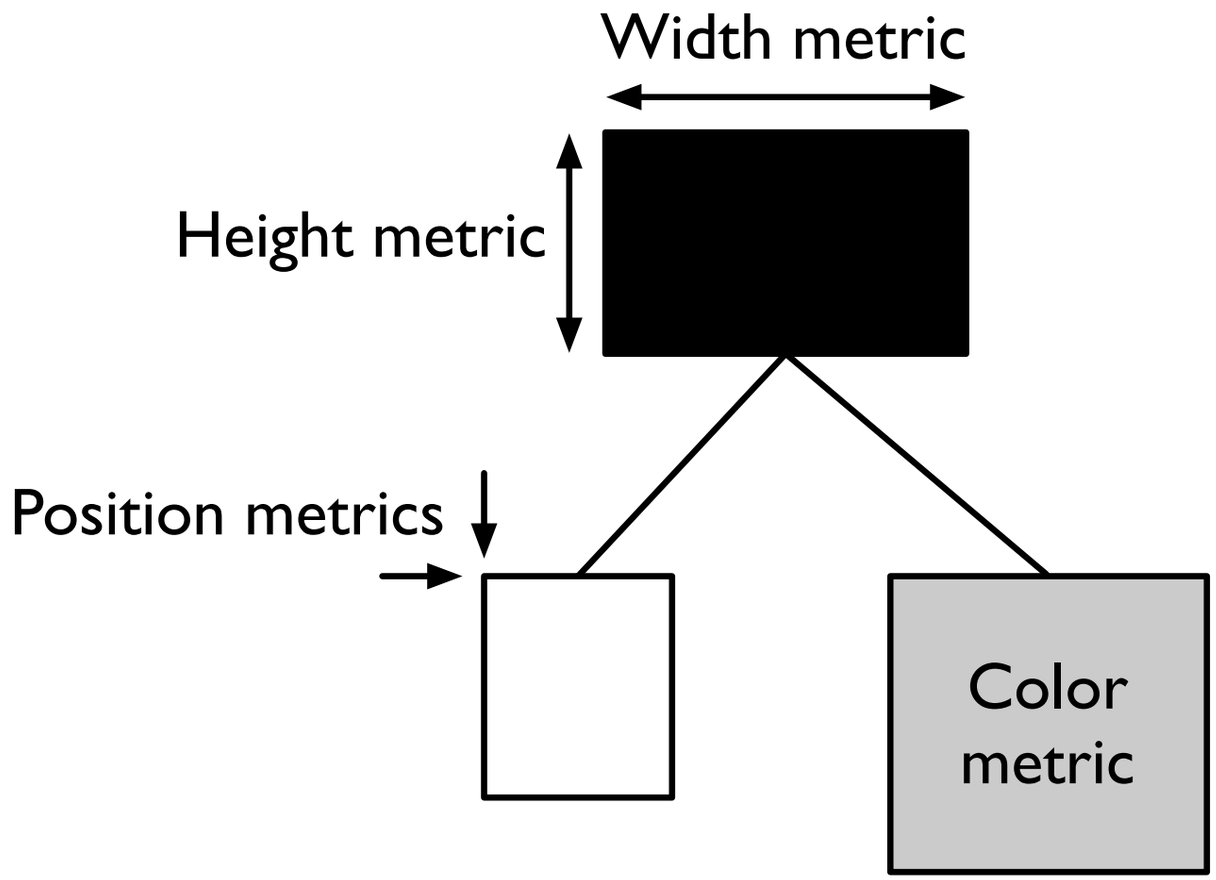




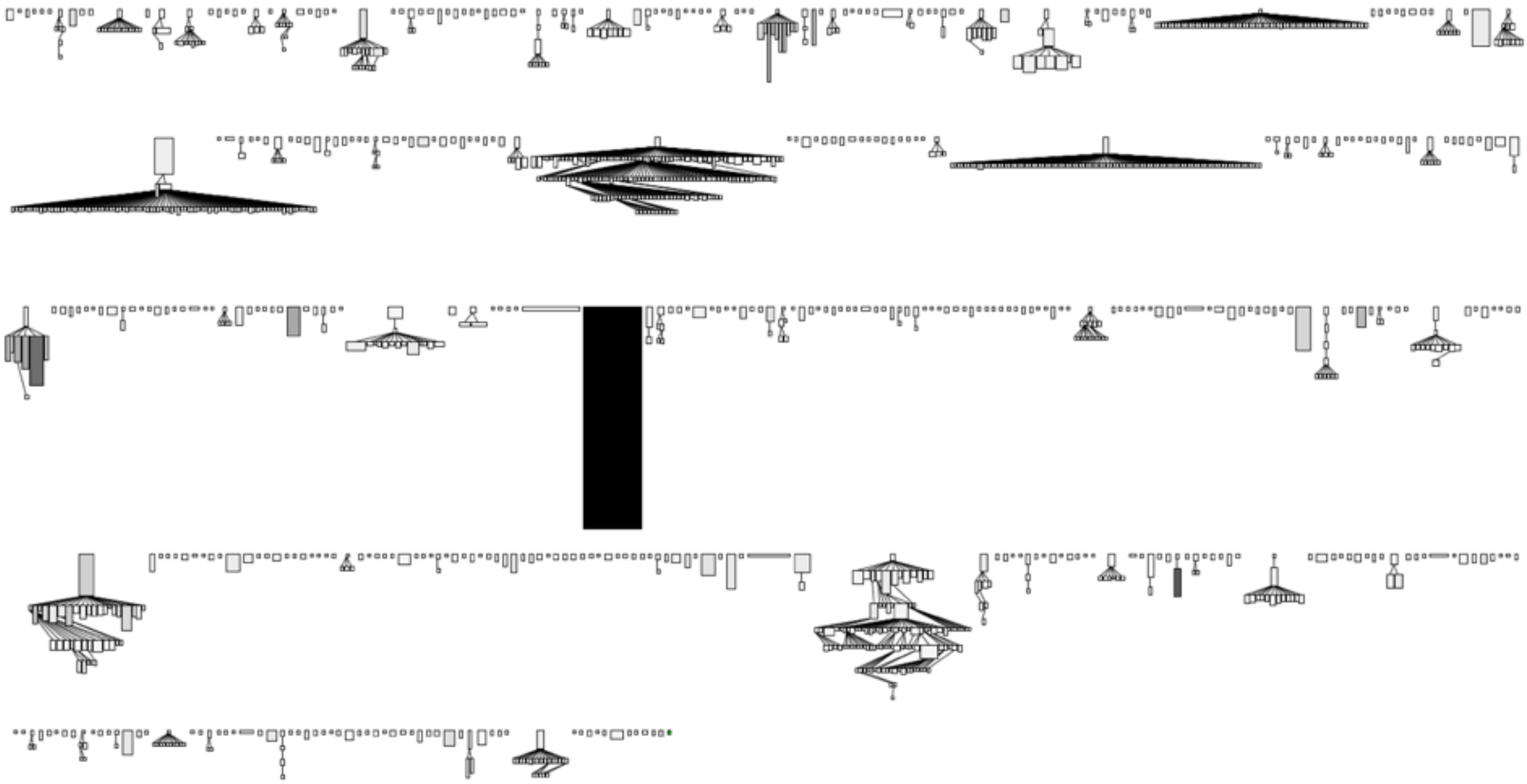
ArgoUML

Polymetric Views

- Use of pre-attentive processing features
  - Size
  - Color
  - Connectedness
- Implemented in…
  - Mondrian, Roassal,
  - XRay



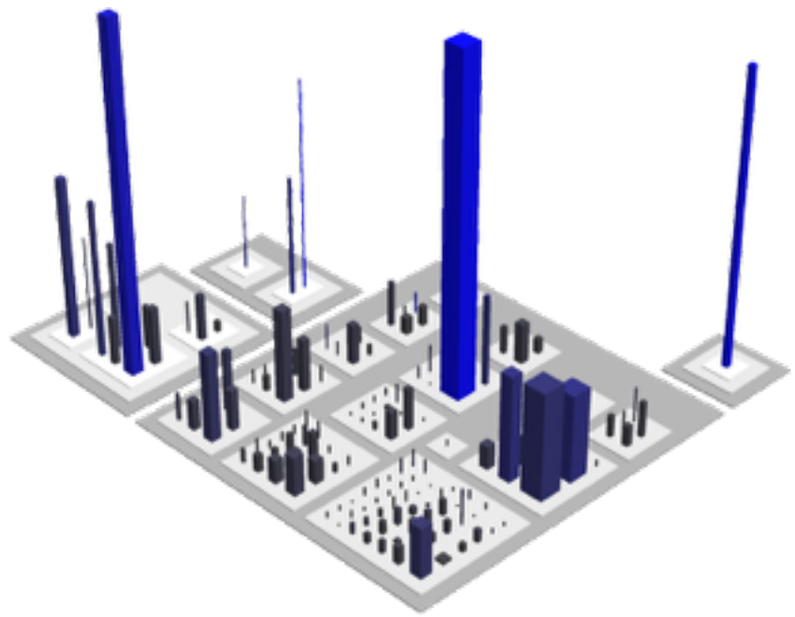
Providing an overview of inheritance



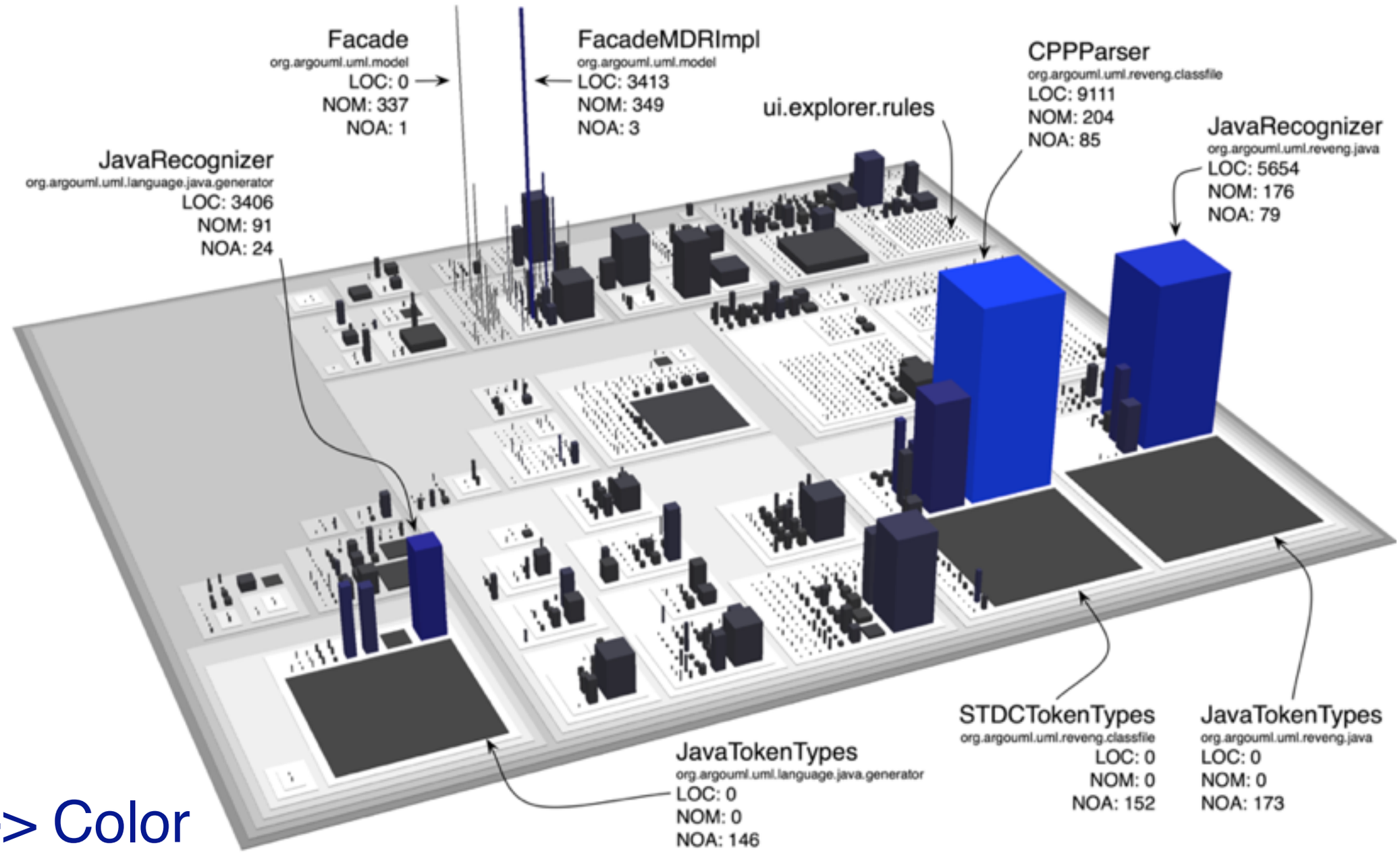
ArgoUML

3D Polymetric Views

- Use of pre-attentive processing features of
  - Size
  - Color
  - 3D spatial locality
- Implemented in…
  - CodeCity (and clones)

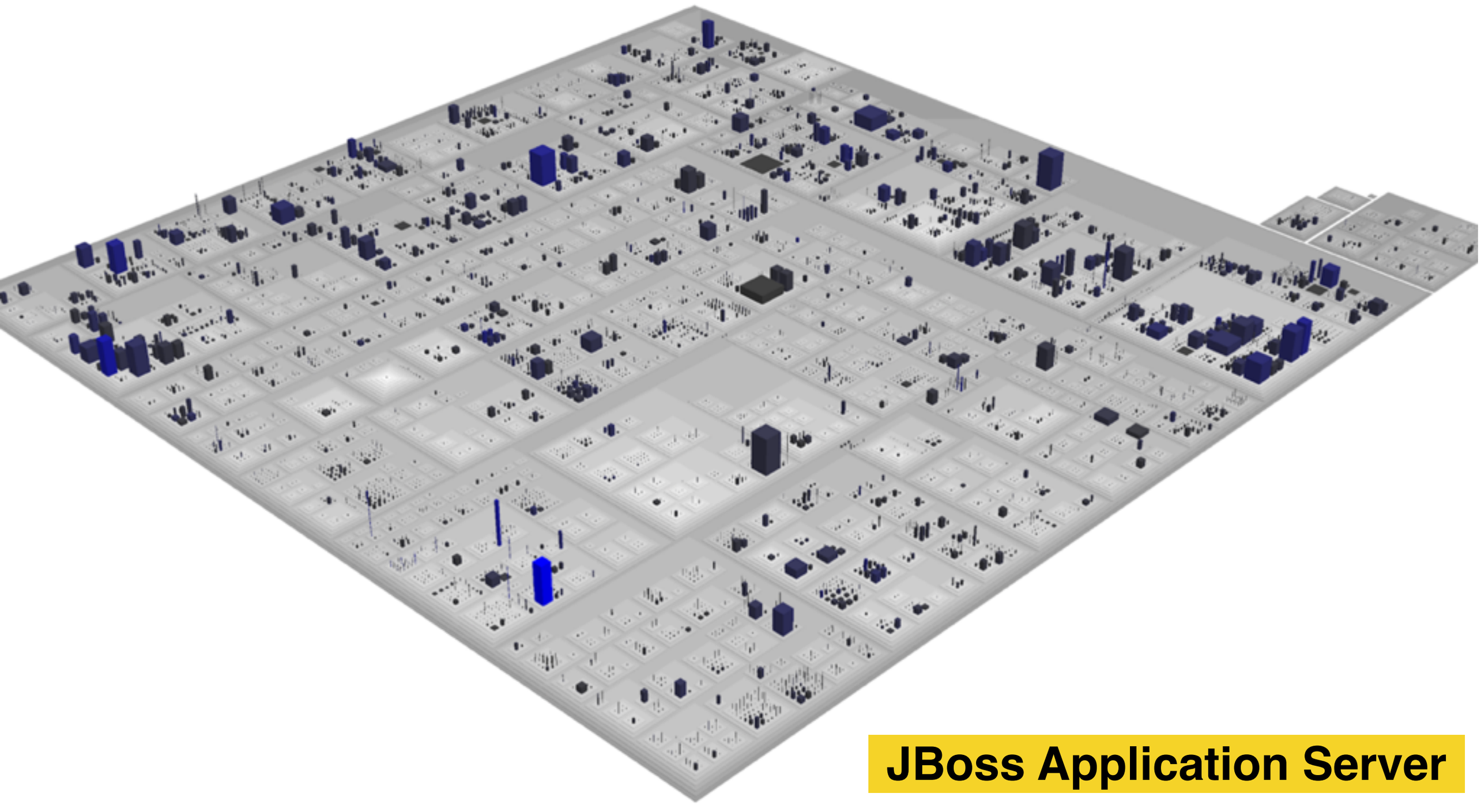


Detecting outliers



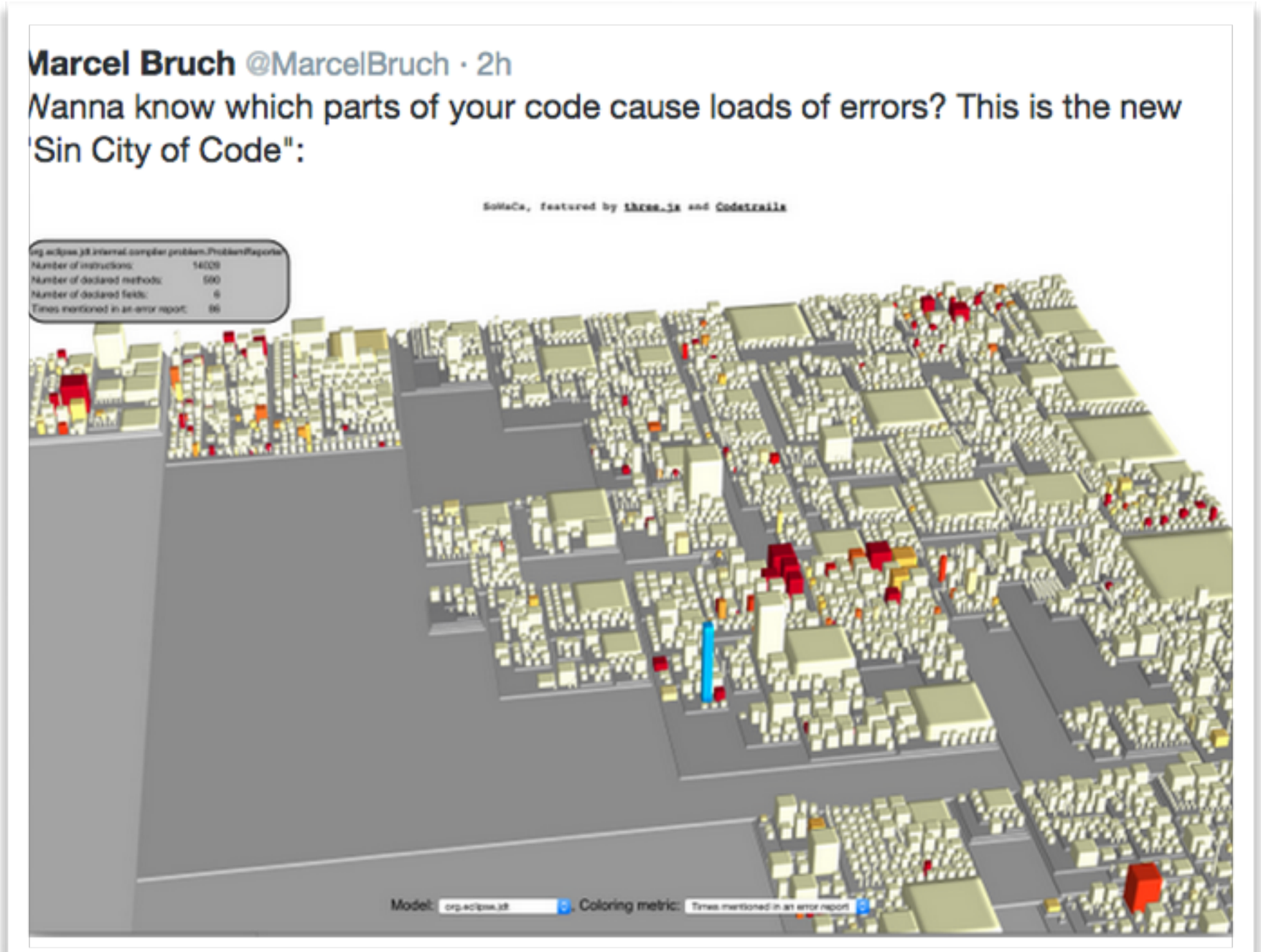
LOC -> Color
NOM -> Height
NOA -> Area

ArgoUML



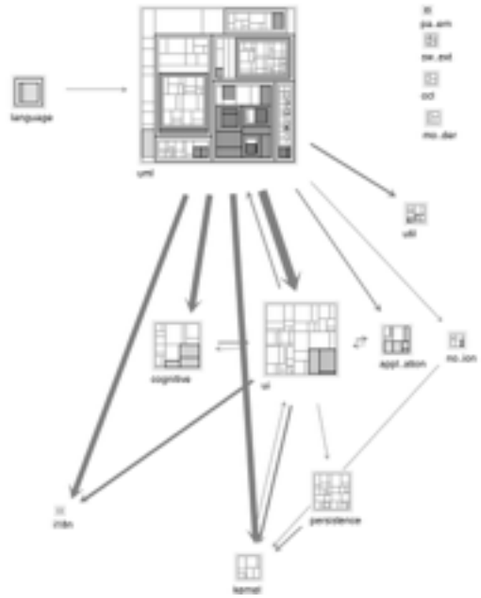
JBoss Application Server

Communicating the locality of problems

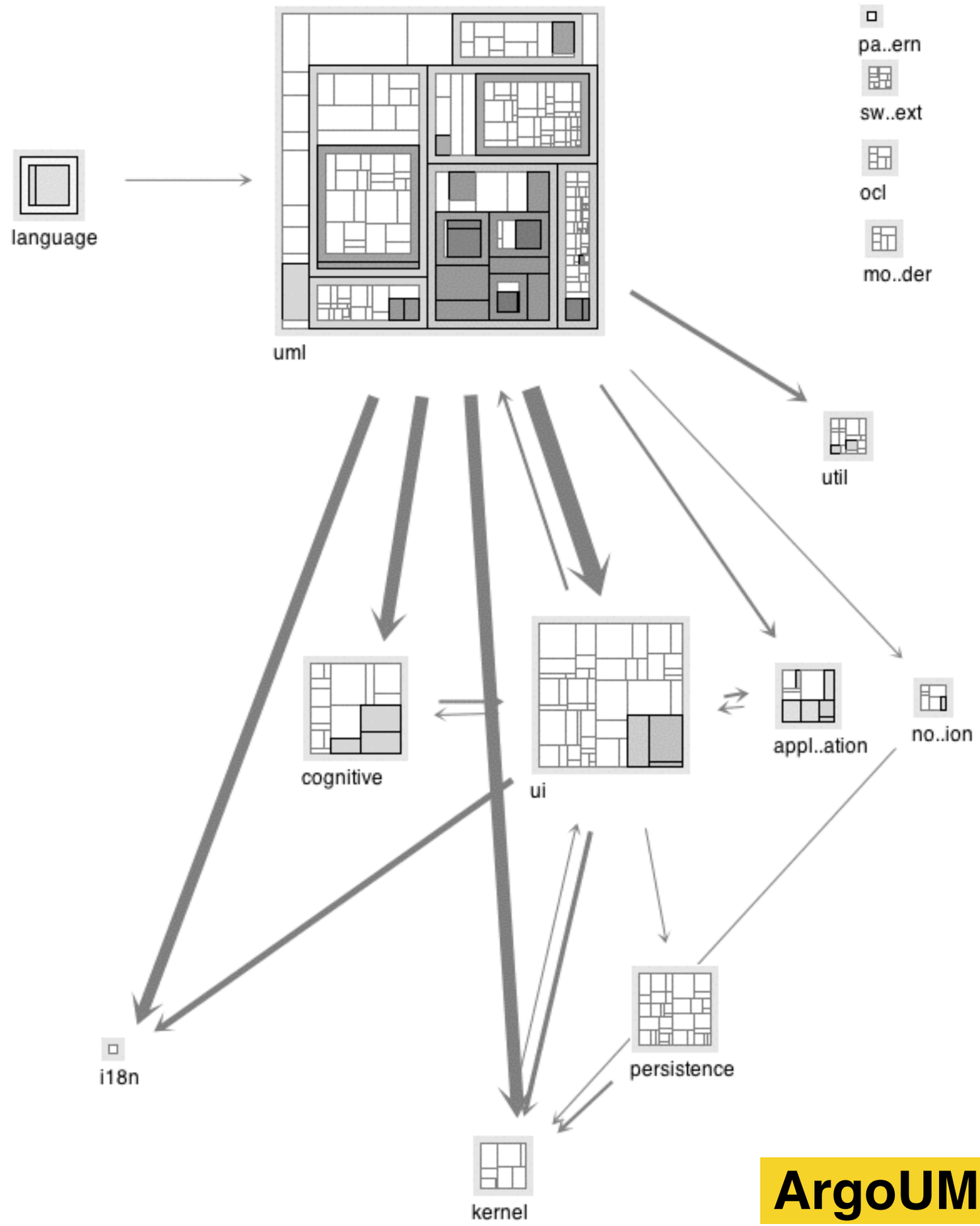


Hierarchical Visualization

- Use of pre-attentive processing features of
  - Size
  - Spatial locality
  - Connectedness
  - Color
- Implemented in…
  - Softwarenaut
  - Rigi, Shrimp, etc.



An overview of the dependencies between the various parts of the system



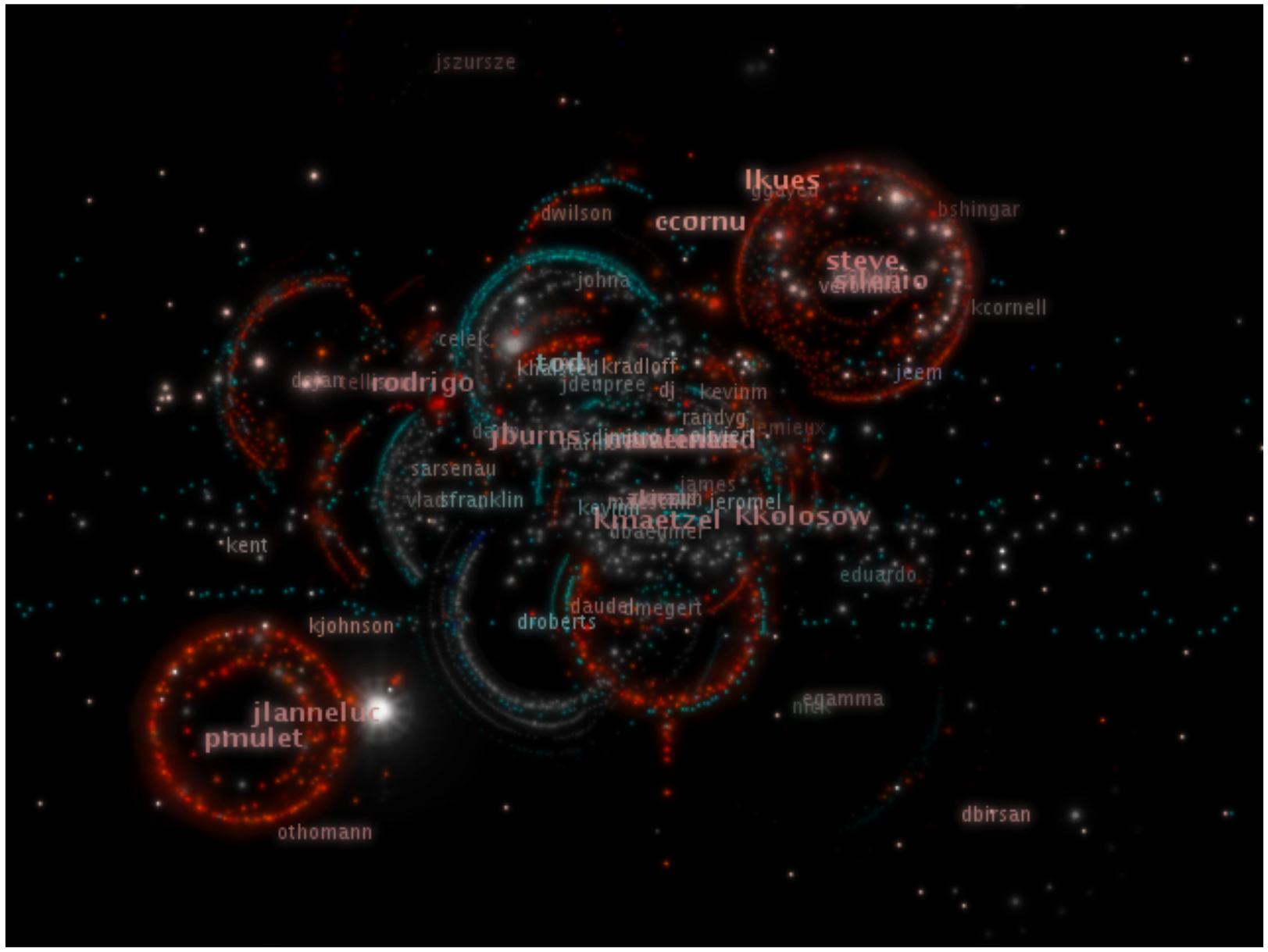
Structure — Summary

- Visualized Aspects
  - Inheritance
  - Containment
  - Dependencies
- Techniques
  - Polymeric Views
  - 3D Polymeric Views
  - Hierarchical Visualization
  - Space filling techniques
- Challenges
  - Displaying both structure and containment

Roadmap

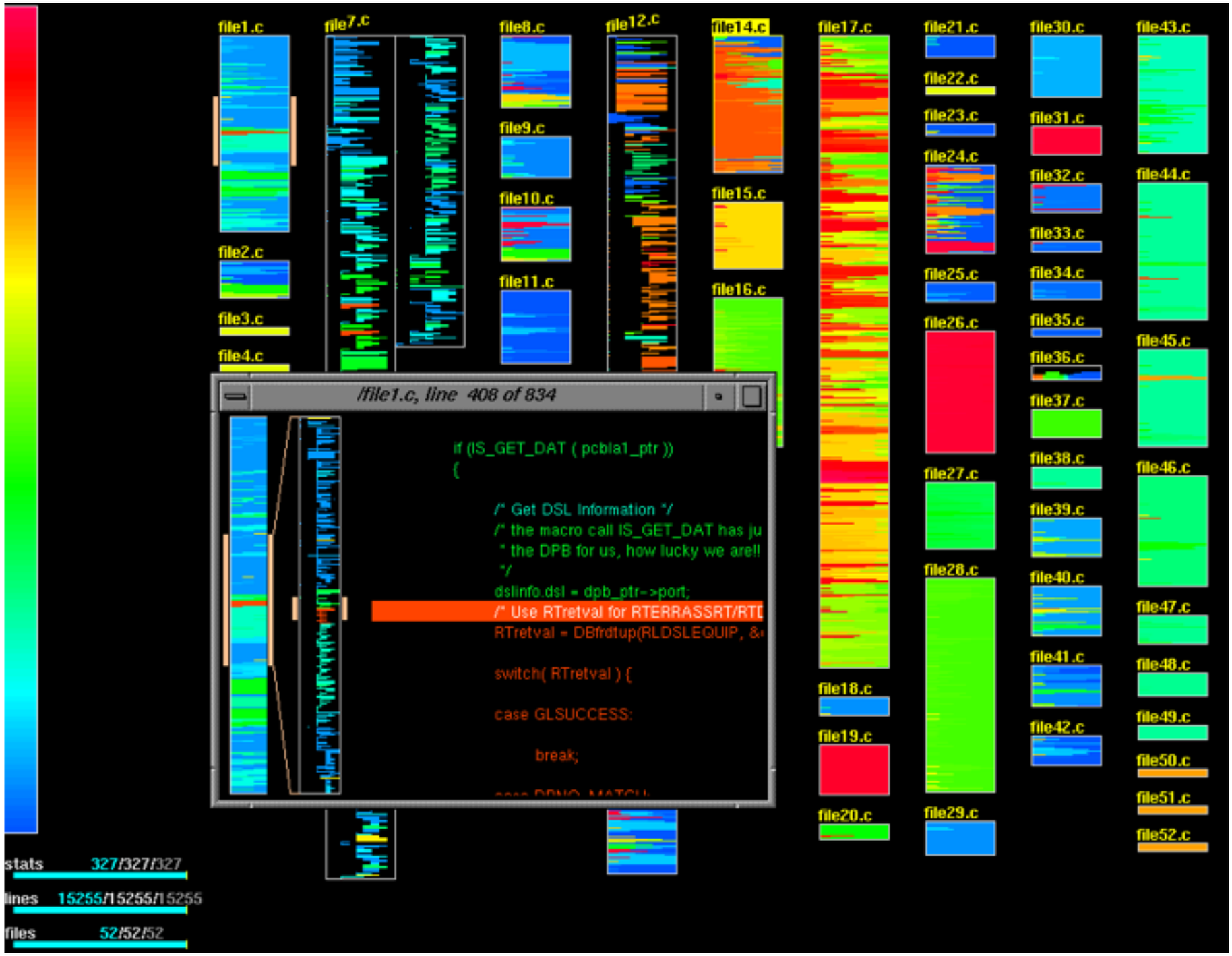
- Information Visualization
- Designing Visualizations
- **Software Visualization**
  - Structure
  - **Evolution**
  - Behavior

Mapping evolution on time

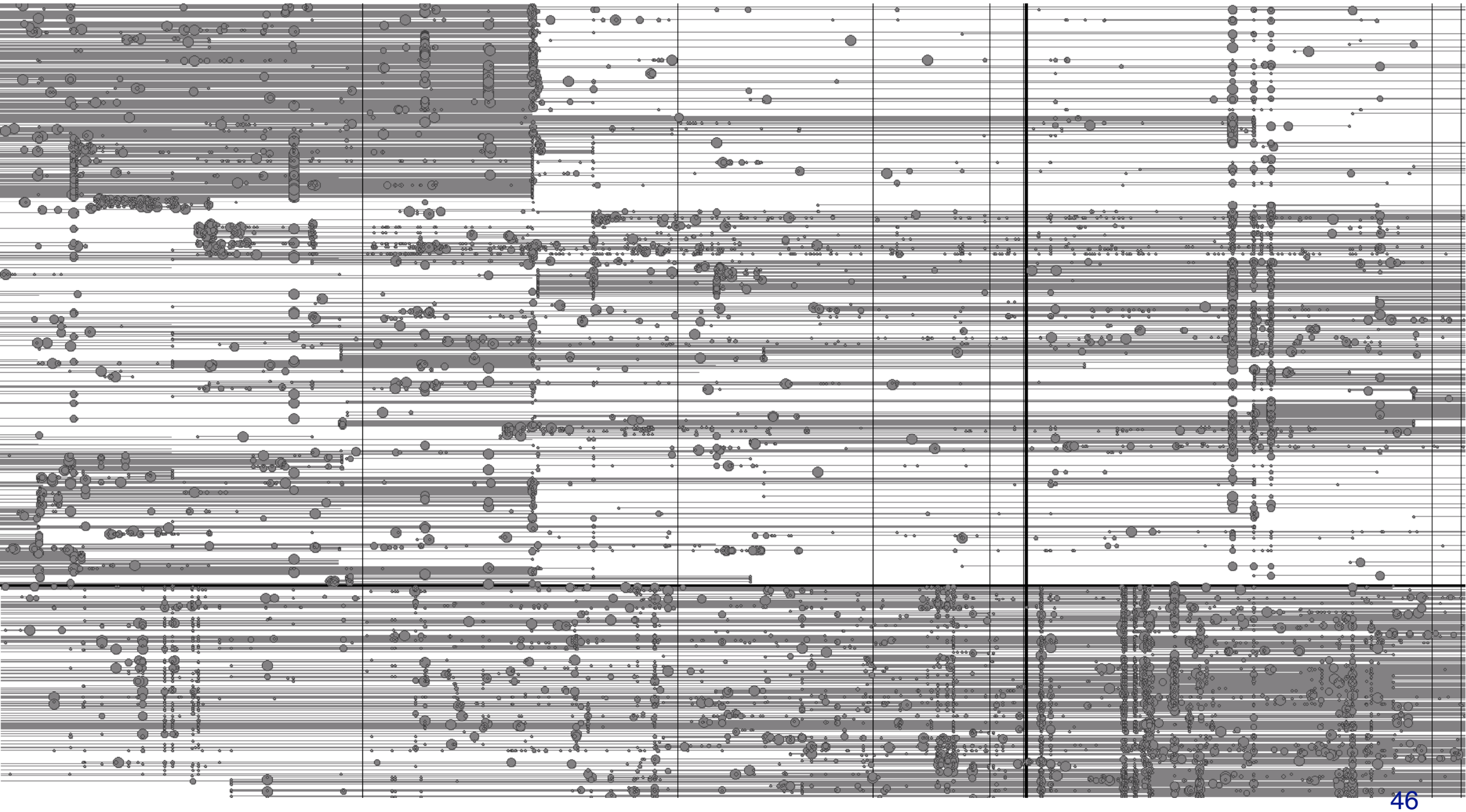


Mapping evolution on color

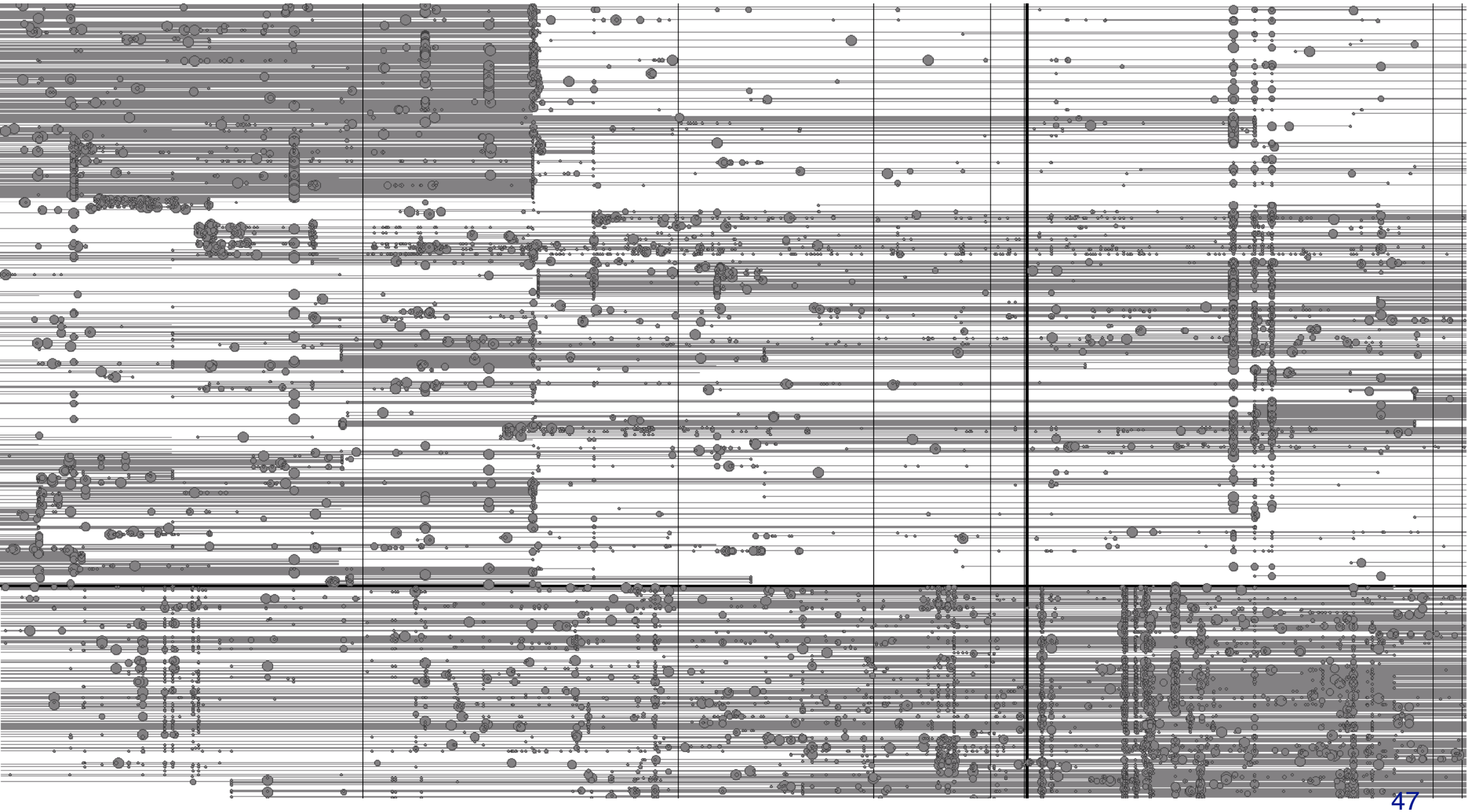
The Seesoft system maps **each line of code into a thin row**. The color of each row indicates a statistic of interest, e.g., **red rows are those most recently changed**, and **blue are those least recently changed**



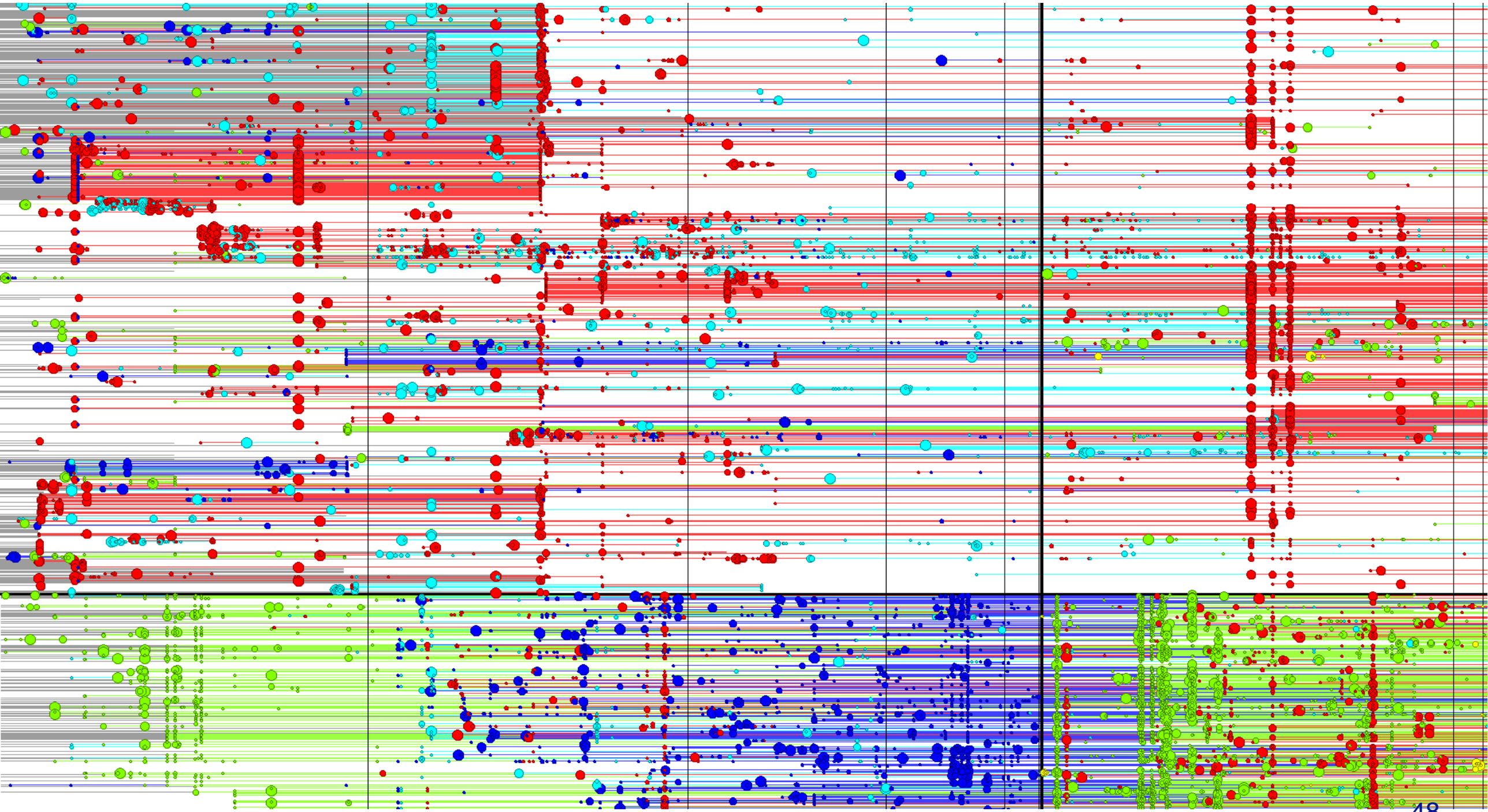
Mapping evolution on space (the x-axis)

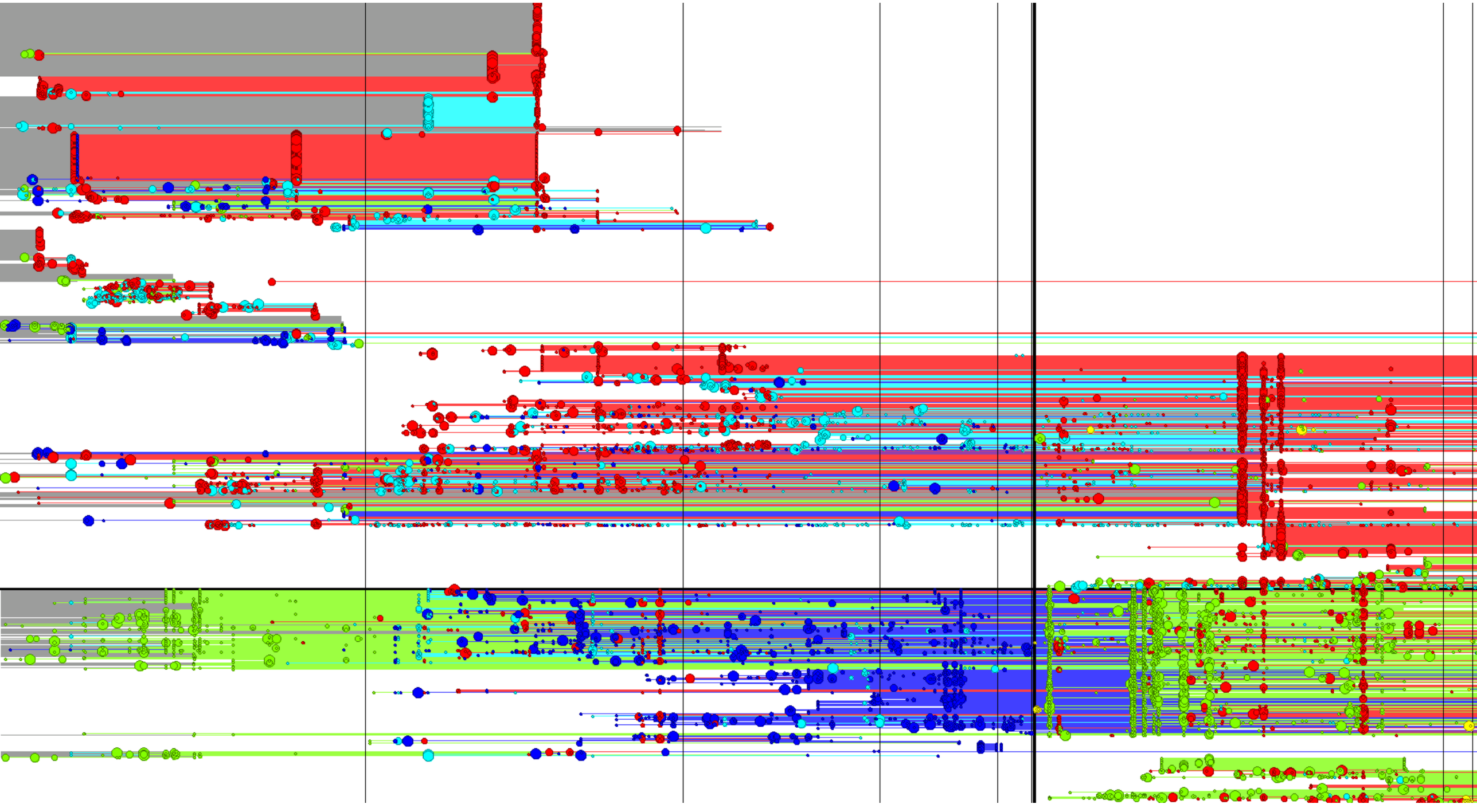


Y-axis represents individual files sorted alphabetically



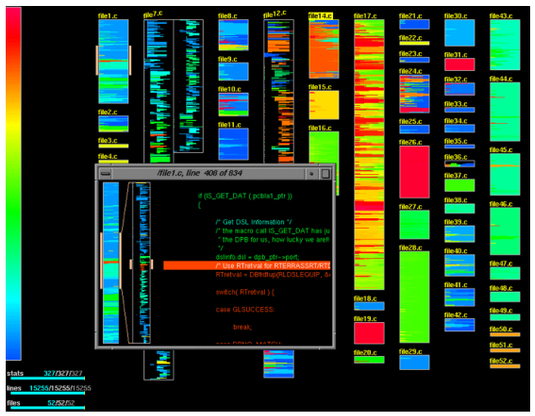
Map authors on colors and kill alphabetical order



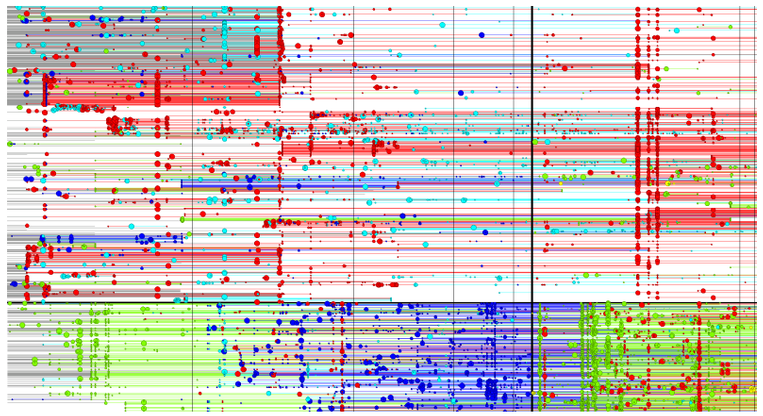


Girba etal,

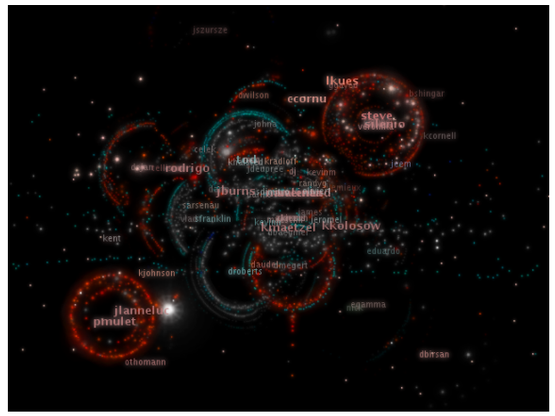
System evolution can be mapped on



color



space

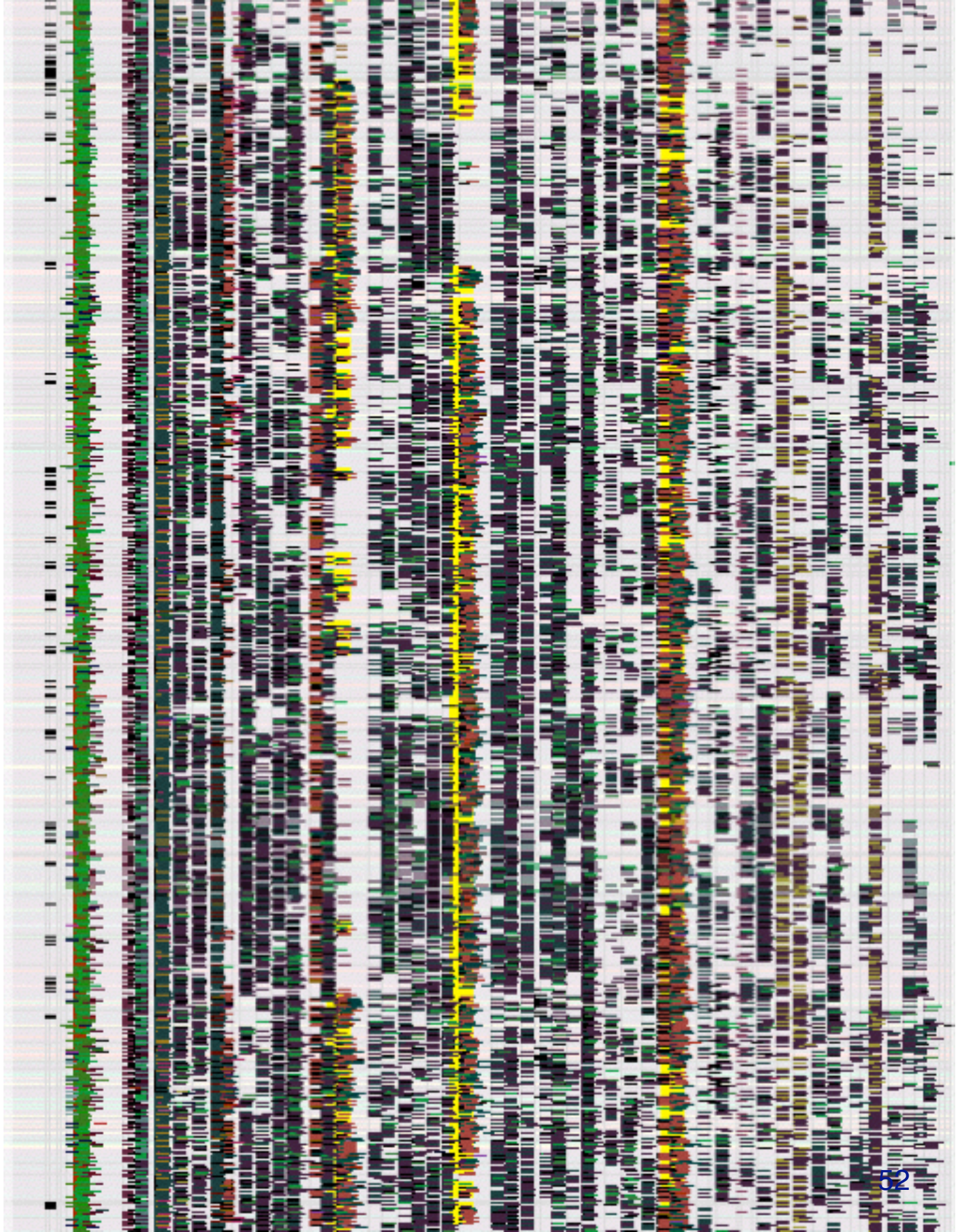


time

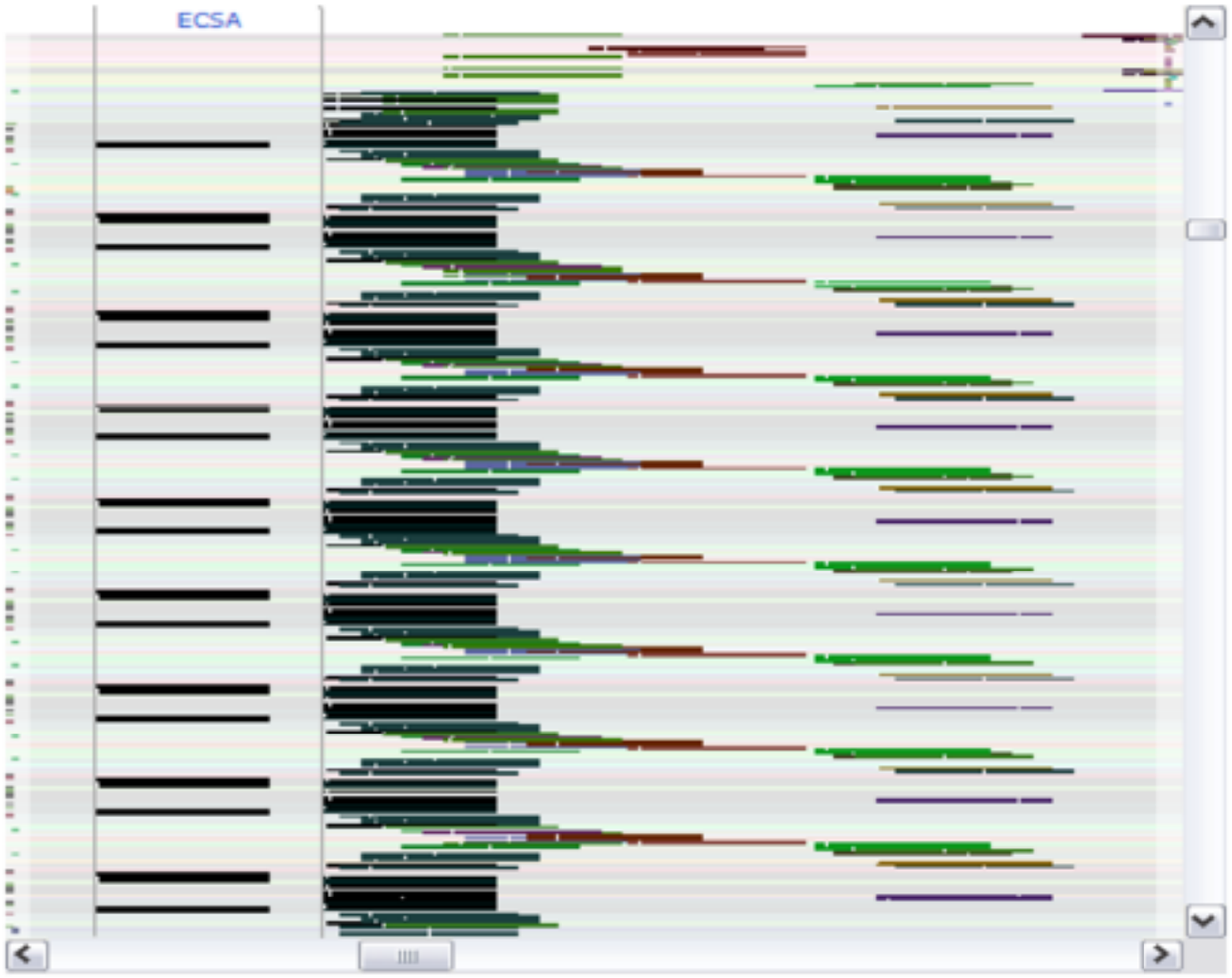
Roadmap

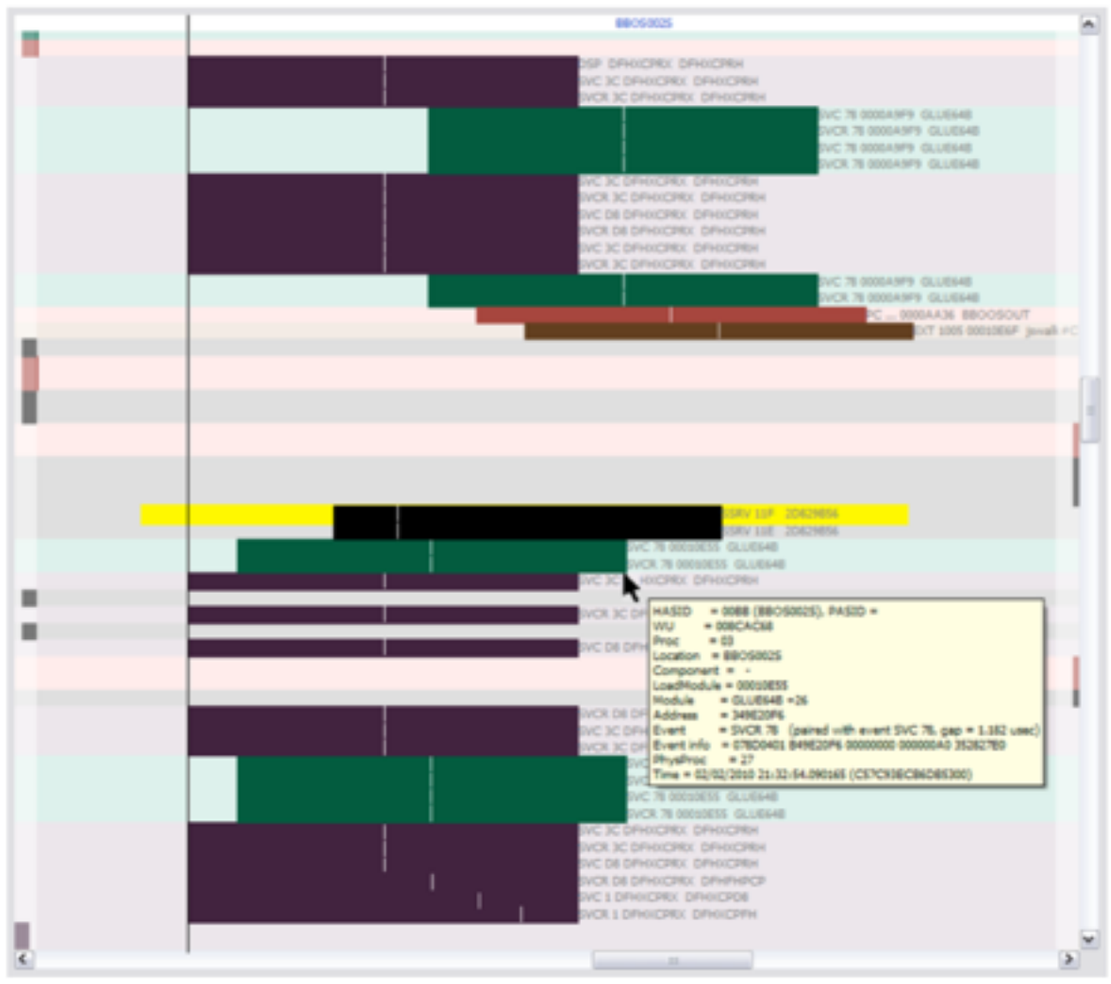
- Information Visualization
- Designing Visualizations
- **Software Visualization**
  - Structure
  - Evolution
  - **Behavior**

Zinsight visualization is targeted at analyzing large event traces



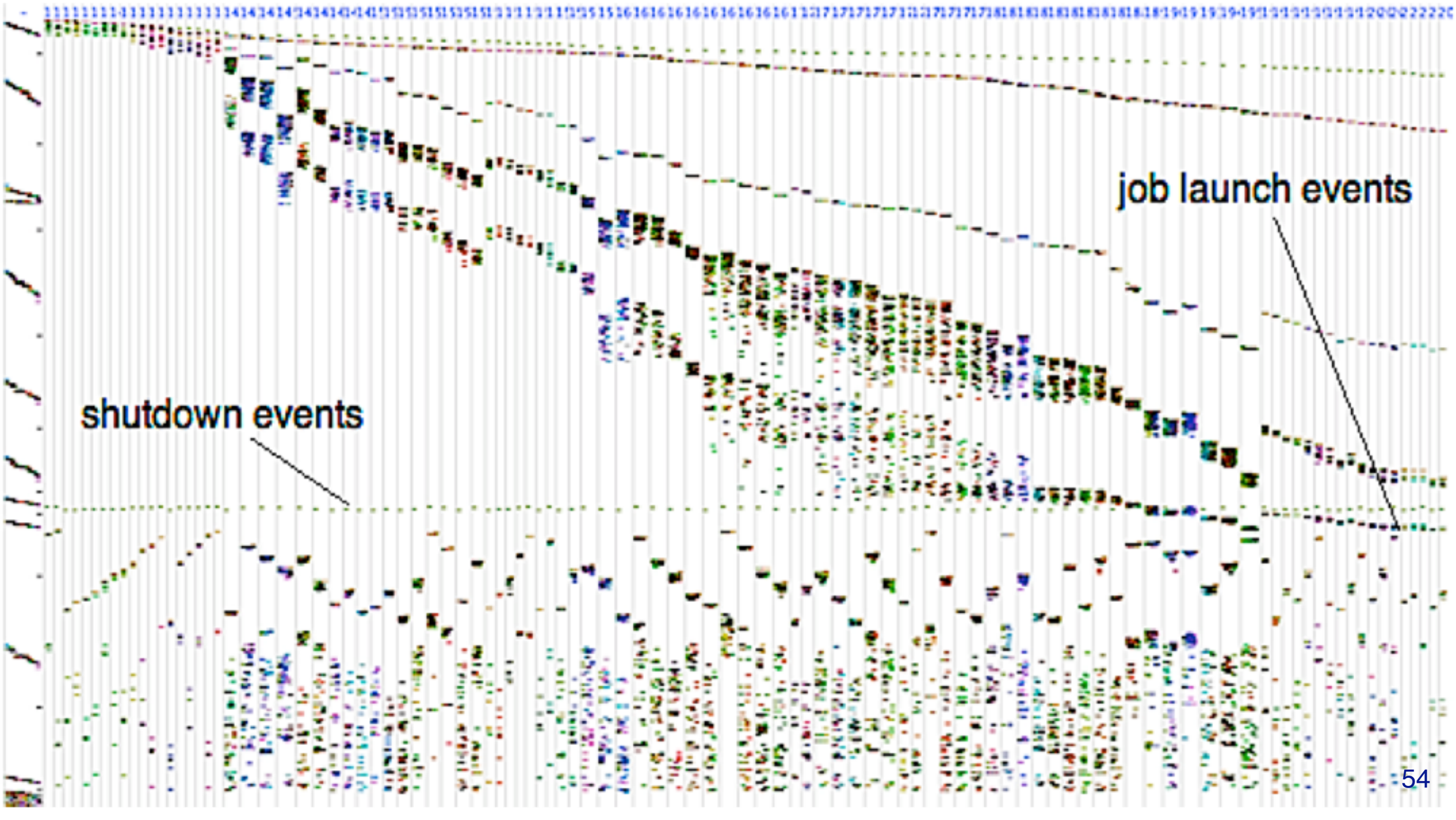
Massively reliant on visual pattern recognition and interactivity

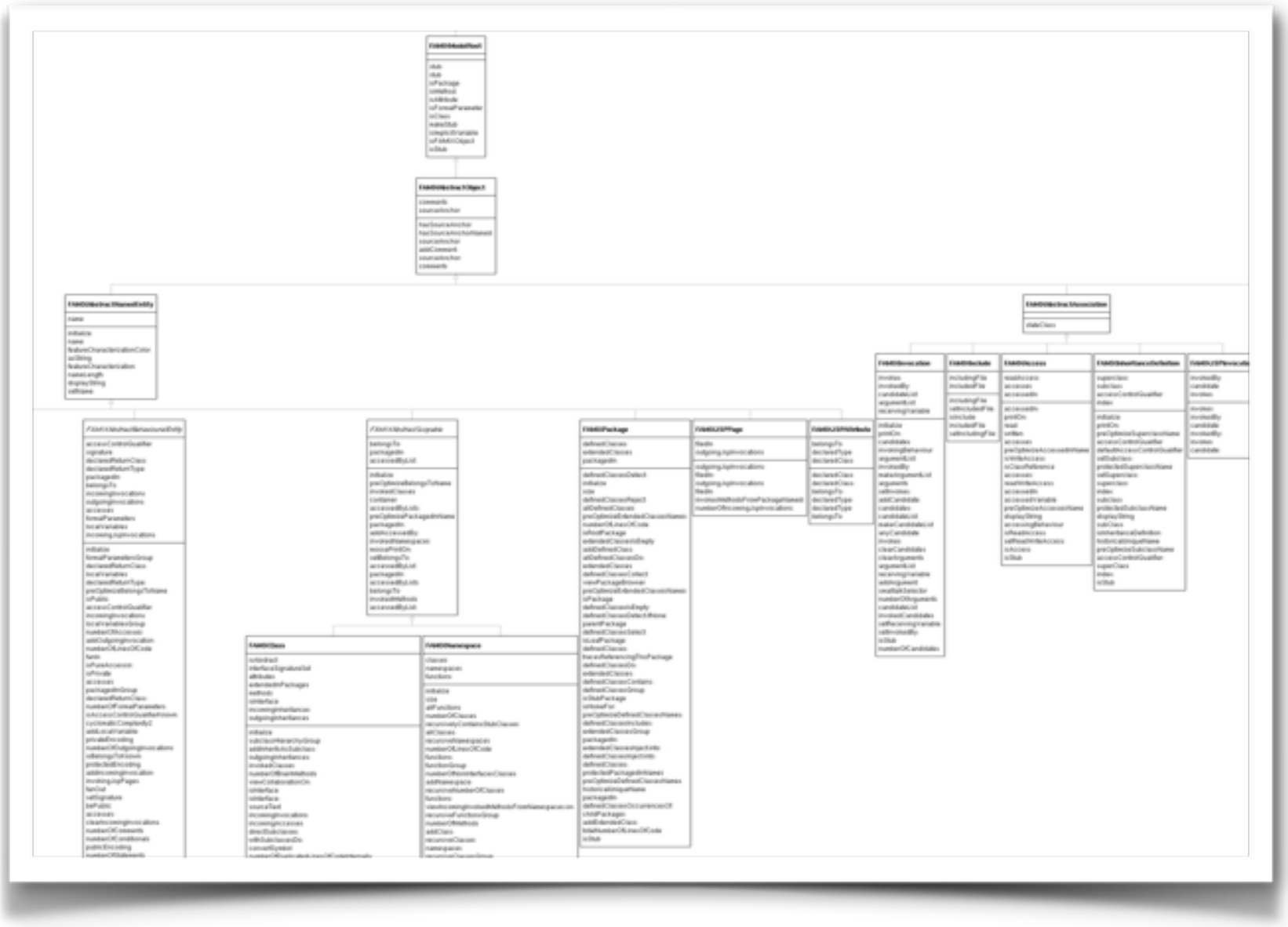




Semantic Zooming

Visual detection of bugs





Ceci n'est pas une visualization.

What you should know

- The laws of Gestalt psychology
- What is information visualization good for
- Which aspects of software can be visualized?
- Which techniques are used in visualizing software structure?
- On what visualization features can we map evolution?
- What kinds of problems can be solved with software visualization?