Dr. Pattabiraman. V

VIT University – Chennai Campus

pattabiraman.v@vit.ac.in

Data Visualisations

Outline

- About Data Visualization
- Data modeling and Visualization Architecture
- Connecting to Data
- Building basic views
- Data manipulations and Calculated fields
- Creating a reports
- Creating a Tableau Dashboard

Visualized data analysis

VIT UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)
VELLORE ▪ CHENNAI
www.vit.ac.in

67 x 89 = ?

```
    67
  x 89
--------
   603
+ 536
--------
  5963
```

Visualization

Representing

information (data)

as

computer graphics.

Scientific, Engineering and Information Visualization

Scientific Visualization: Scientific Data

Engineering Visualization: Measurement Data

Information Visualization: Abstract Data

Visualization Analysis & Design

VIT UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)
VELLORE ▪ CHENNAI
www.vit.ac.in

- Defining visualization
  - **Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.**

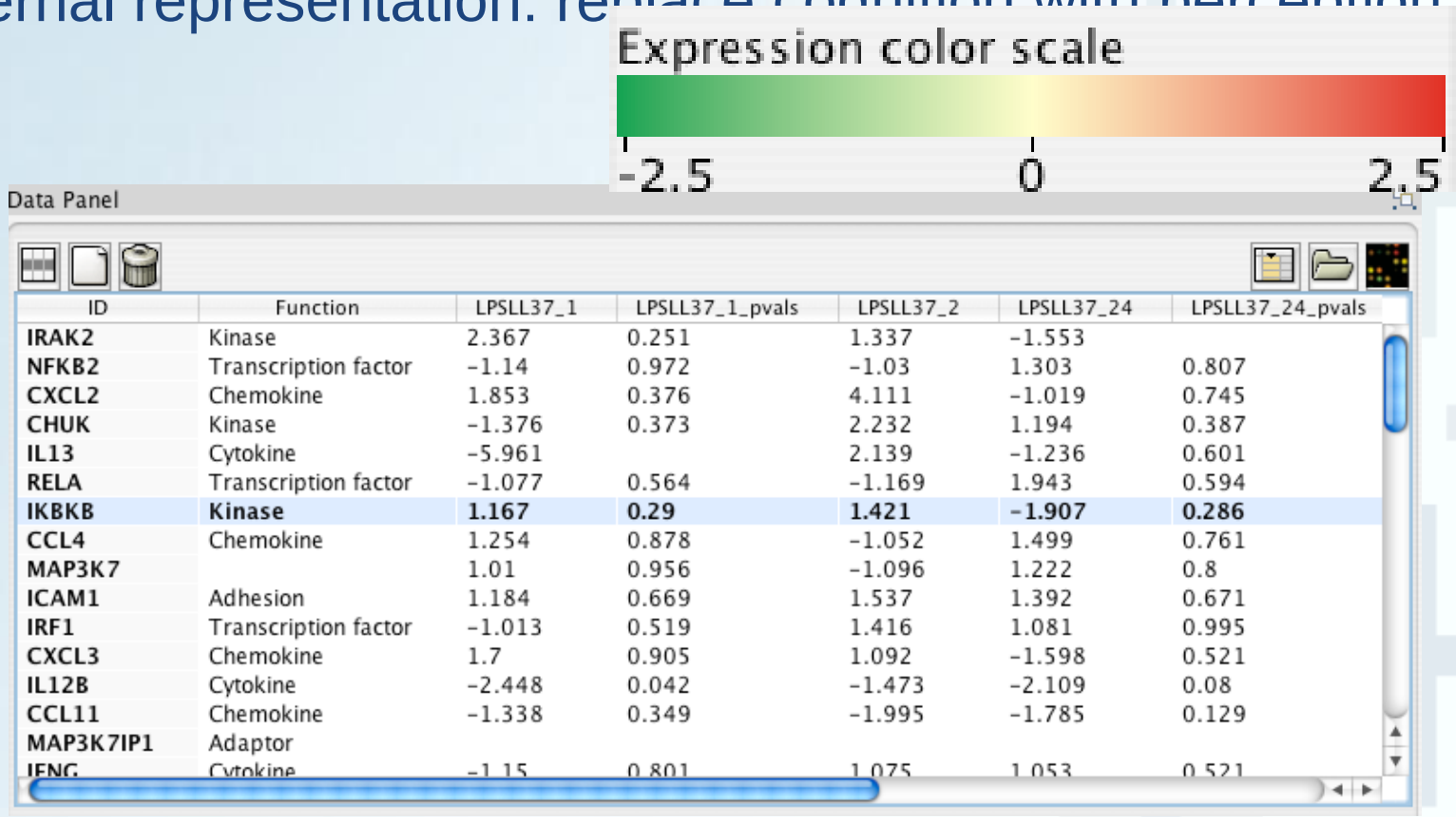
Datasets and People

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational decision-making methods.

- long-term use for end users (e.g. exploratory analysis of scientific data)
- presentation of known results
- stepping stone to better understanding of requirements before developing models
- help developers of automatic solution refine/debug, determine parameters
- help end users of automatic solutions verify, build trust

Visual Representation

external representation: replace cognition with perception



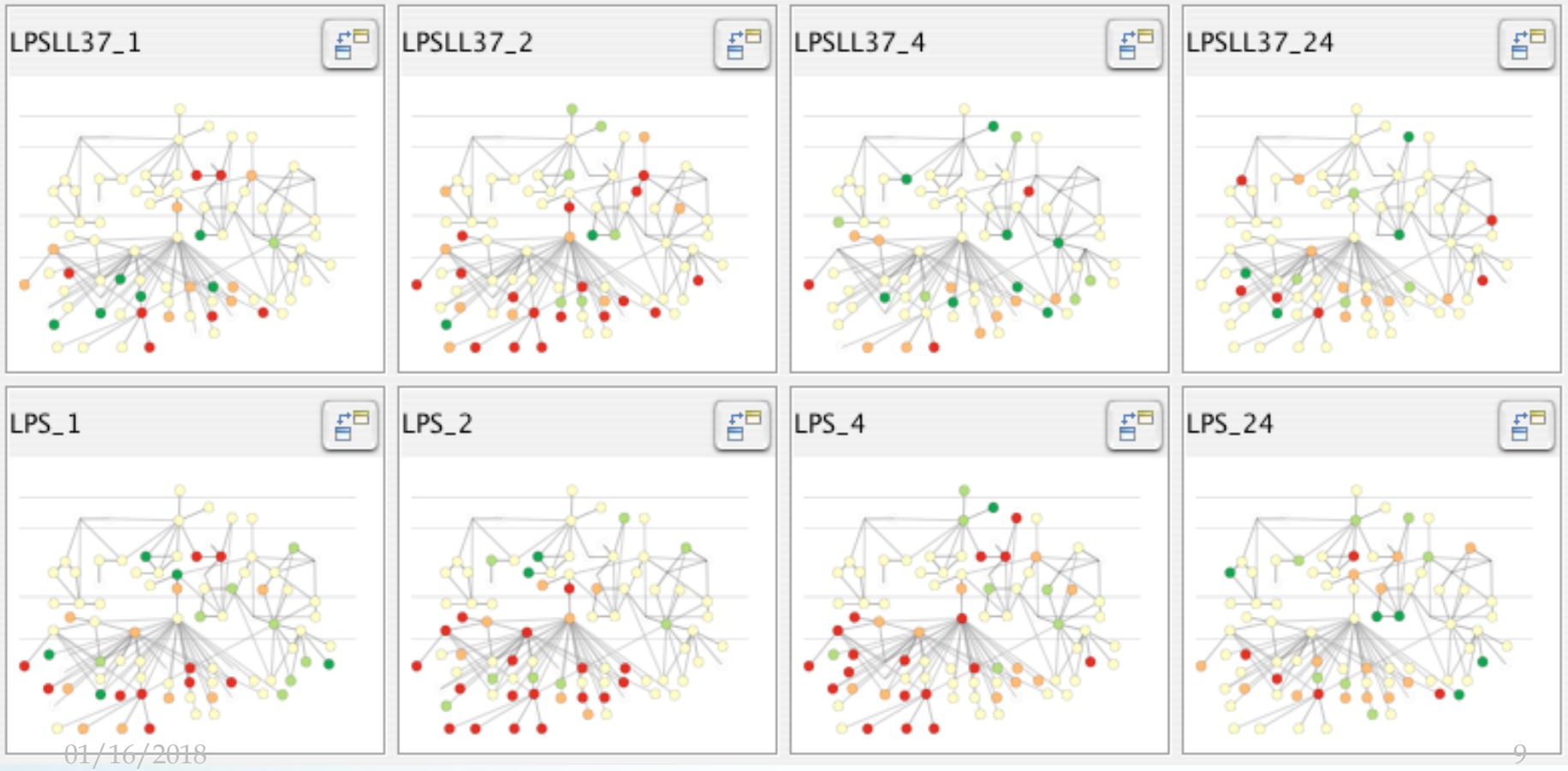
ID	Function	LPSLL37_1	LPSLL37_1_pvals	LPSLL37_2	LPSLL37_24	LPSLL37_24_pvals
IRAK2	Kinase	2.367	0.251	1.337	-1.553	
NFKB2	Transcription factor	-1.14	0.972	-1.03	1.303	0.807
CXCL2	Chemokine	1.853	0.376	4.111	-1.019	0.745
CHUK	Kinase	-1.376	0.373	2.232	1.194	0.387
IL13	Cytokine	-5.961		2.139	-1.236	0.601
RELA	Transcription factor	-1.077	0.564	-1.169	1.943	0.594
IKBKB	**Kinase**	**1.167**	**0.29**	**1.421**	**-1.907**	**0.286**
CCL4	Chemokine	1.254	0.878	-1.052	1.499	0.761
MAP3K7		1.01	0.956	-1.096	1.222	0.8
ICAM1	Adhesion	1.184	0.669	1.537	1.392	0.671
IRF1	Transcription factor	-1.013	0.519	1.416	1.081	0.995
CXCL3	Chemokine	1.7	0.905	1.092	-1.598	0.521
IL12B	Cytokine	-2.448	0.042	-1.473	-2.109	0.08
CCL11	Chemokine	-1.338	0.349	-1.995	-1.785	0.129
MAP3K7IP1	Adaptor					
IFNG	Cytokine	-1.15	0.801	1.075	1.053	0.521

[Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context. Barsky, Munzner, Gardy, and Kincaid. IEEE TVCG (Proc. InfoVis) 14(6):1253-1260, 2008.]

Visual Representation

VIT UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)
VELLORE ▪ CHENNAI
www.vit.ac.in

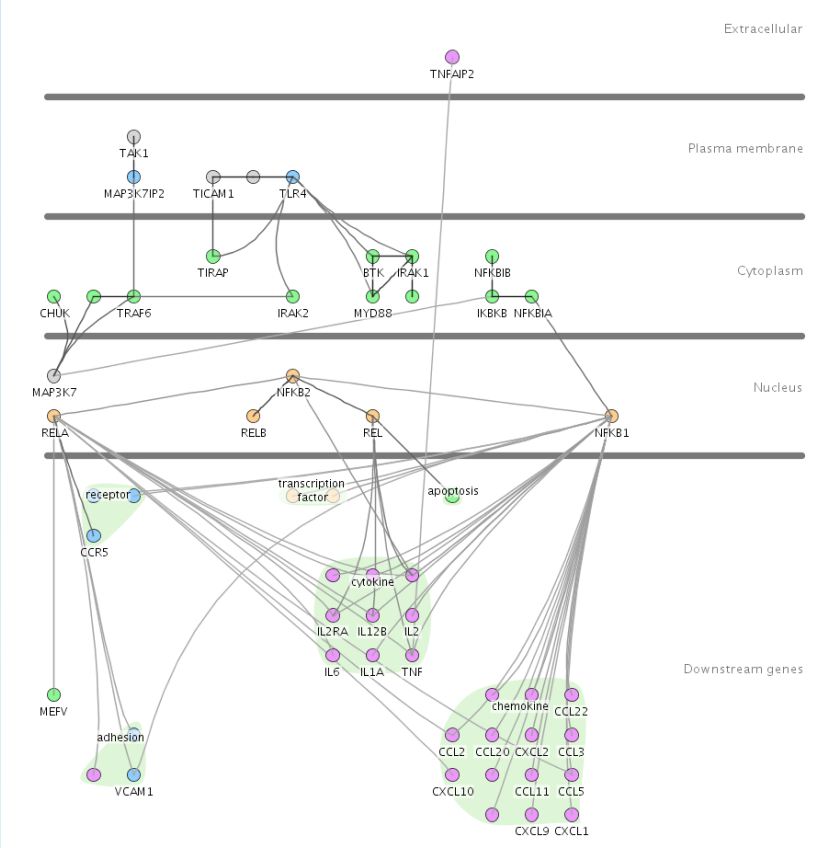
external representation: replace cognition with perception

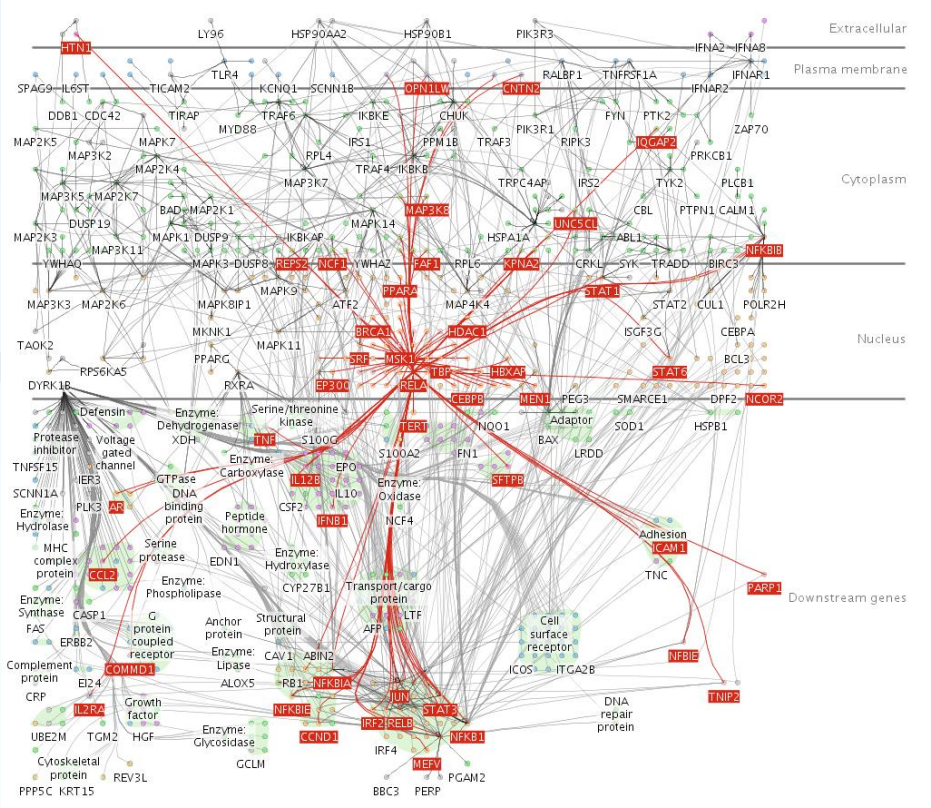


Computer Based Visualization

VIT UNIVERSITY
(Estd. u/s 3 of UGC Act 1956)
VELLORE ■ CHENNAI
www.vit.ac.in

beyond human patience: scale to large datasets, support interactivity





Analysis: What, why, and how

- what is shown?
  - **data** abstraction
- why is the user looking at it?
  - **task** abstraction
- how is it shown?
  - **idiom**: visual encoding and interaction
- abstract vocabulary avoids domain-specific terms
  - translation process iterative, tricky
- what-why-how analysis framework as scaffold to think systematically about design space

What?

Datasets | **Attributes**

Data Types

→ Items → Attributes → Links → Positions → Grids

Data and Dataset Types

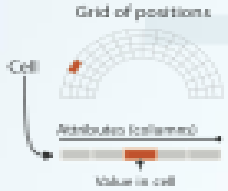
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

→ Tables

→ Networks

→ Fields (Continuous)



→ *Multidimensional Table*

→ *Trees*

→ Geometry (Spatial)

Dataset Availability

→ Static

→ Dynamic

Attribute Types

→ Categorical

→ Ordered

→ *Ordinal*

→ *Quantitative*

Ordering Direction

→ Sequential

→ Diverging

→ Cyclic

Why?

Actions

Analyze

- Consume
  - Discover
  - Present
  - Enjoy
- Produce
  - Annotate
  - Record
  - Derive

Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

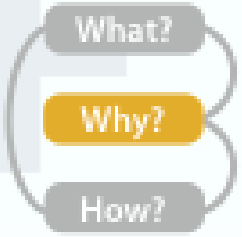
- One
- Many
  - Distribution
  - Dependency
  - Correlation
  - Similarity
  - Extremes

Network Data

- Topology
  - Paths

Spatial Data

- Shape



What?

- Tree

Why?

- Actions
  - Present
  - Locate
  - Identify

- Targets
  - Path between two nodes

How?

- SpaceTree
  - Encode
  - Navigate
  - Select
  - Filter
  - Aggregate

- TreeJuxtaposer
  - Encode
  - Navigate
  - Select
  - Arrange

Encode

- **Arrange**
  - → Express
  - → Separate
  - → Order
  - → Align
  - → Use

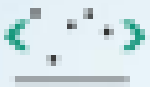
- **Map**
  from categorical and ordered attributes
  - → Color
    - → *Hue*
    - → *Saturation*
    - → *Luminance*
  - → Size, Angle, Curvature, ...
  - → Shape
  - → Motion
    Direction, Rate, Frequency, ...

Manipulate

- Change

- Select

- Navigate

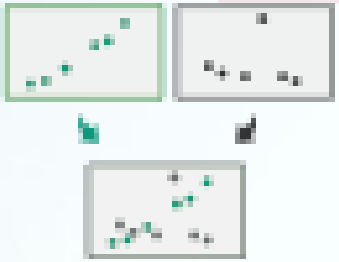


Facet

- Juxtapose

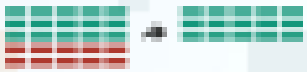
- Partition

- Superimpose



Reduce

- Filter

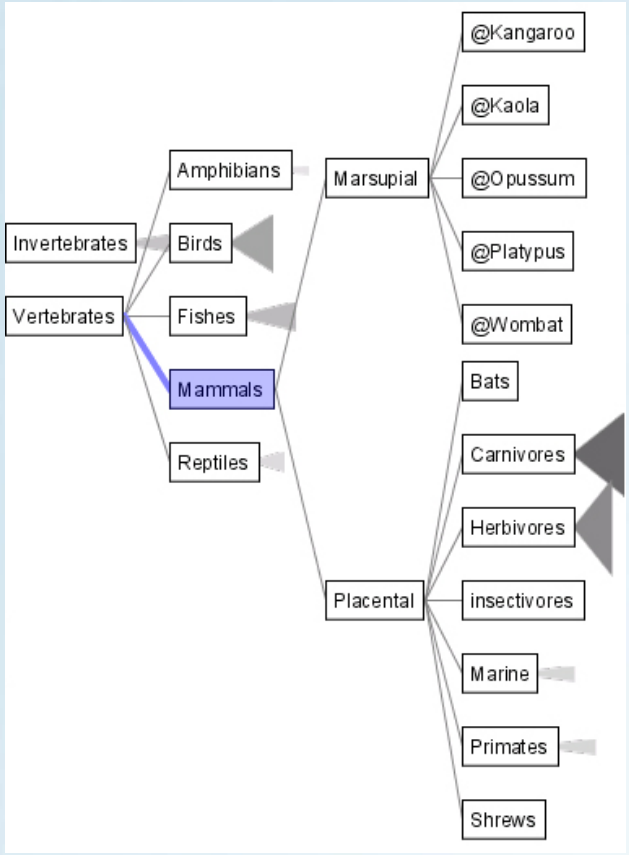


- Aggregate

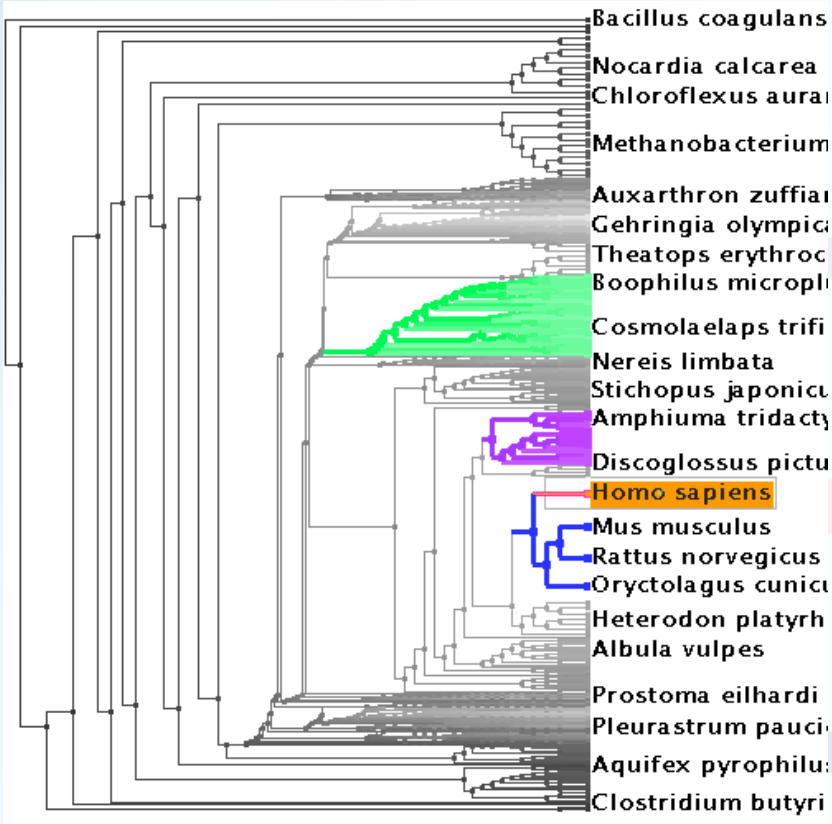
- Embed

Analysis example: Compare idioms

SpaceTree



TreeJuxtaposer

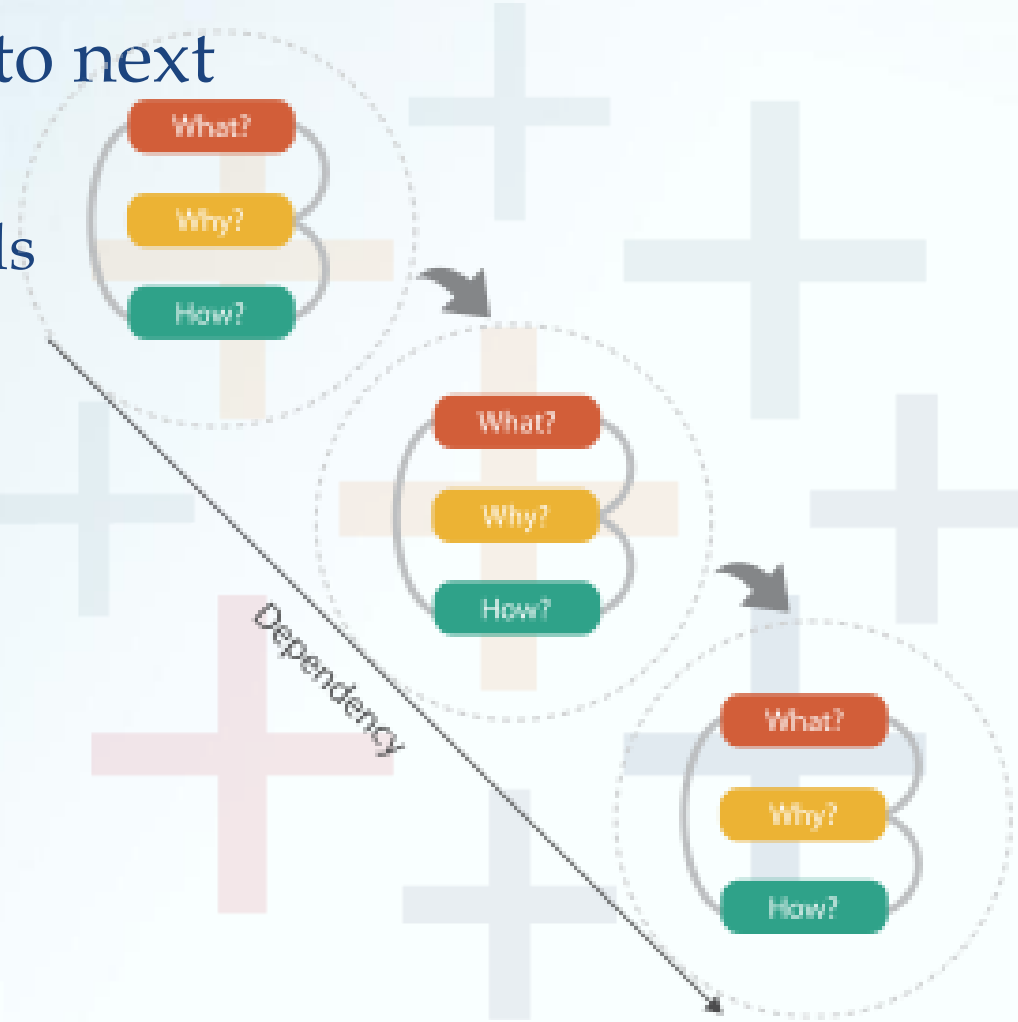


[SpaceTree: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. Grosjean, Plaisant, and Bederson. Proc. InfoVis 2002, p 57–64.]

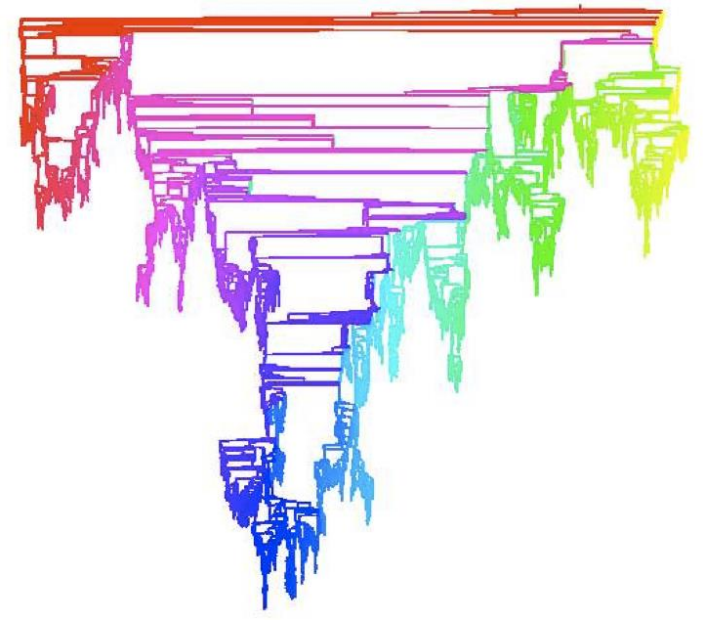
[TreeJuxtaposer: Scalable Tree Comparison Using Focus+Context With Guaranteed Visibility. ACM Trans. on Graphics (Proc. SIGGRAPH) 22:453– 462, 2003.]

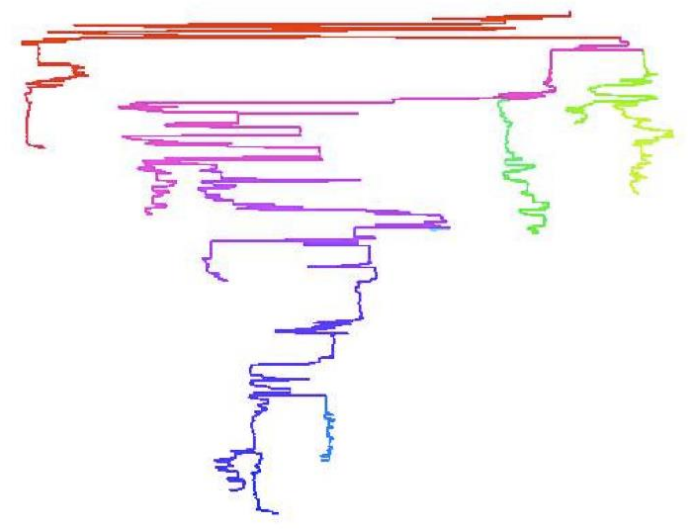
Chained sequences

- output of one is input to next
  - express dependencies
  - separate means from ends



Analysis example: Derive one attribute





- Strahler number
  - centrality metric for trees/networks
  - derived quantitative attribute
  - draw top 5K of 500K for good skeleton

Task 1

In
Tree

Out
Quantitative attribute on nodes

What?	Why?
In Tree	Derive
Out Quantitative attribute on nodes	

Task 2

In
Tree

+

In
Quantitative attribute on nodes

Out
Filtered Tree
Removed unimportant parts

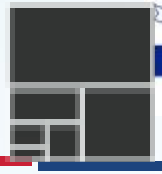
What?	Why?	How?
In Tree	Summarize	Reduce
In Quantitative attribute on nodes	Topology	Filter
Out Filtered Tree		

Definitions: Marks and channels

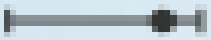
- Marks - geometric primitives
- Channels - control appearance of marks
  - can redundantly code with multiple channels
- interactions
  - point marks only convey position; no area constraints - can be size and shape coded
  - line marks convey position and length
    - can only be size coded in 1D (width)
  - area marks fully constrained
    - cannot be size or shape coded

- Points
- Lines
- Areas

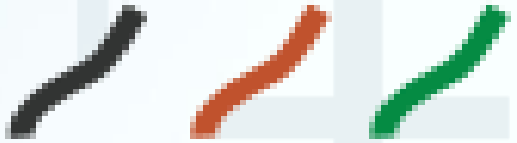


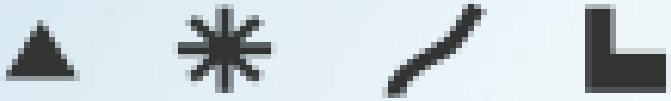
- Position
  - Horizontal
  - Vertical
  - Both



- Color

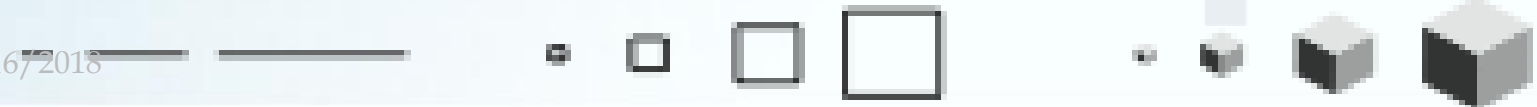


- Shape



- Tilt

- Size
  - Length
  - Area
  - Volume



THANK YOU