

Visualization & Journalism: Four Vignettes

Tamara Munzner

Department of Computer Science

University of British Columbia

Computation and Journalism Symposium 2016

October 2016, Stanford CA

<http://www.cs.ubc.ca/~tmm/talks.html#cj16>

[@tamaramunzner](https://twitter.com/tamaramunzner)

Four vignettes

- a tale of two tools created for journalistic use
 - shared frameworks of interdisciplinary methods from my research group
 - thinking about collaboration
 - roles & rewards, for computer scientists & journalists
 - reasoning about visualization design
 - beyond pretty pictures
 - divergent goals & audiences
 - Overview: investigation / exploratory
 - TimeLineCurator: presentation / explanatory
- two cautionary tales with actionable advice
 - lessons we've learned in vis
 - challenges of color
 - difficulties of depth

Visualization (vis) defined & motivated

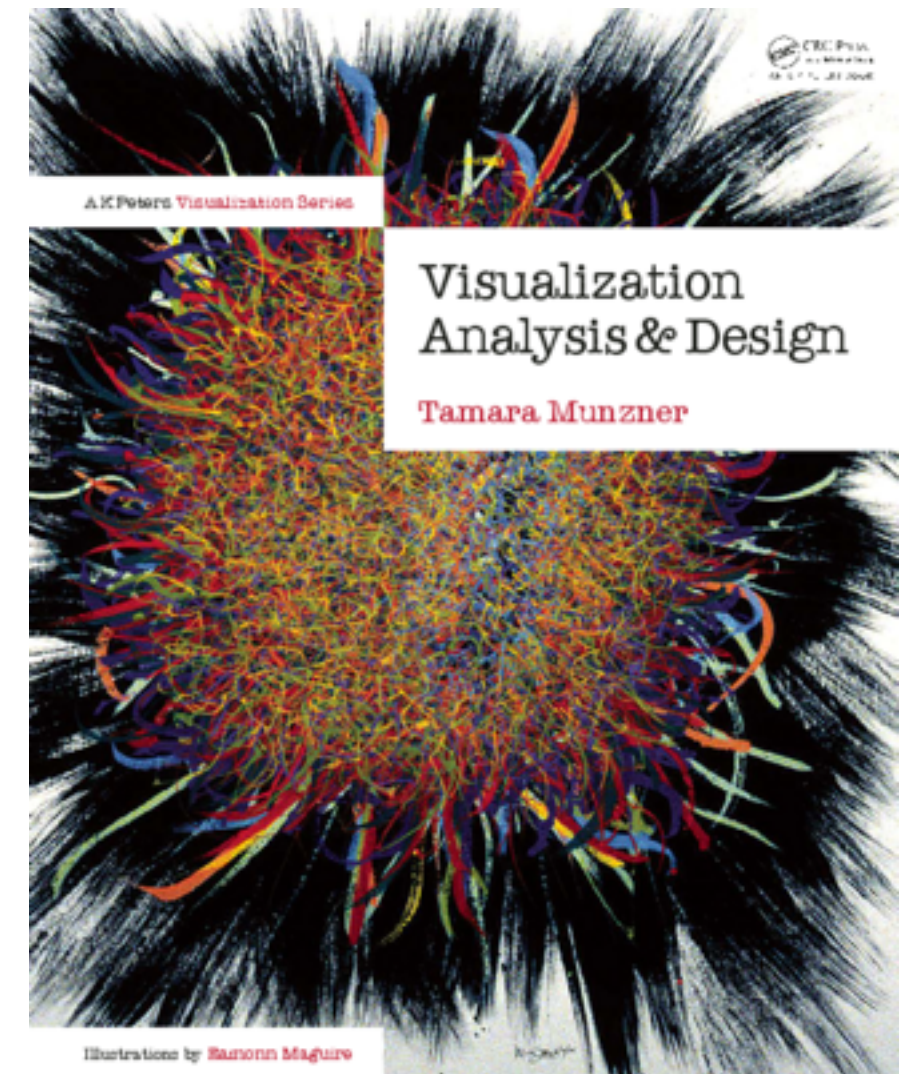
Computer-based visualization systems provide visual representations of datasets designed to help people carry out tasks more effectively.

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

- human in the loop needs the details
 - doesn't know exactly what questions to ask in advance
 - longterm exploratory analysis
 - presentation of known results
 - stepping stone towards automation: refining, trustbuilding
- external representation: perception vs cognition
- intended task, measurable definitions of effectiveness

more at:

Visualization Analysis and Design, Chapter 1.
Munzner. AK Peters Visualization Series, CRC Press, 2014.



Vignette 1: Vis Tool for Investigative Reporting



Matthew Brehmer
@mattbrehmer



Stephen Ingram
@FroweFace



Overview

The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists

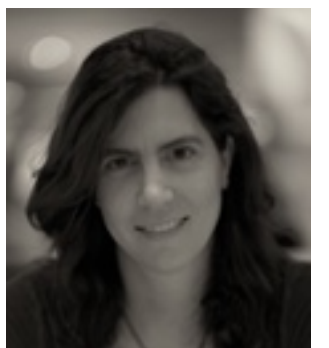
<http://www.cs.ubc.ca/labs/imager/tr/2014/Overview/>

<https://www.overviewdocs.com>

Jonathan Stray
@jonathanstray



Tamara Munzner
@tamaramunzner



Overview: The Design, Adoption, and Analysis of a Visual Document Mining Tool For Investigative Journalists. Brehmer, Ingram, Stray, and, Munzner. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2014)*, 20(12):2271-2280, 2014.

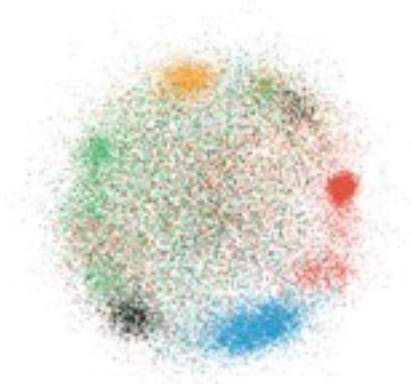
Origin story: WikiLeaks meets Glimmer

- WikiLeaks: hacker-journalist Jonathan Stray analyzing Iraq warlogs
 - one instance of general problem: Too Many Documents
 - conjectured that existing label classification falls short of showing all meaningful structure in data
 - friendly action, criminal incident, ...
 - he had some NLP, needed better vis tools

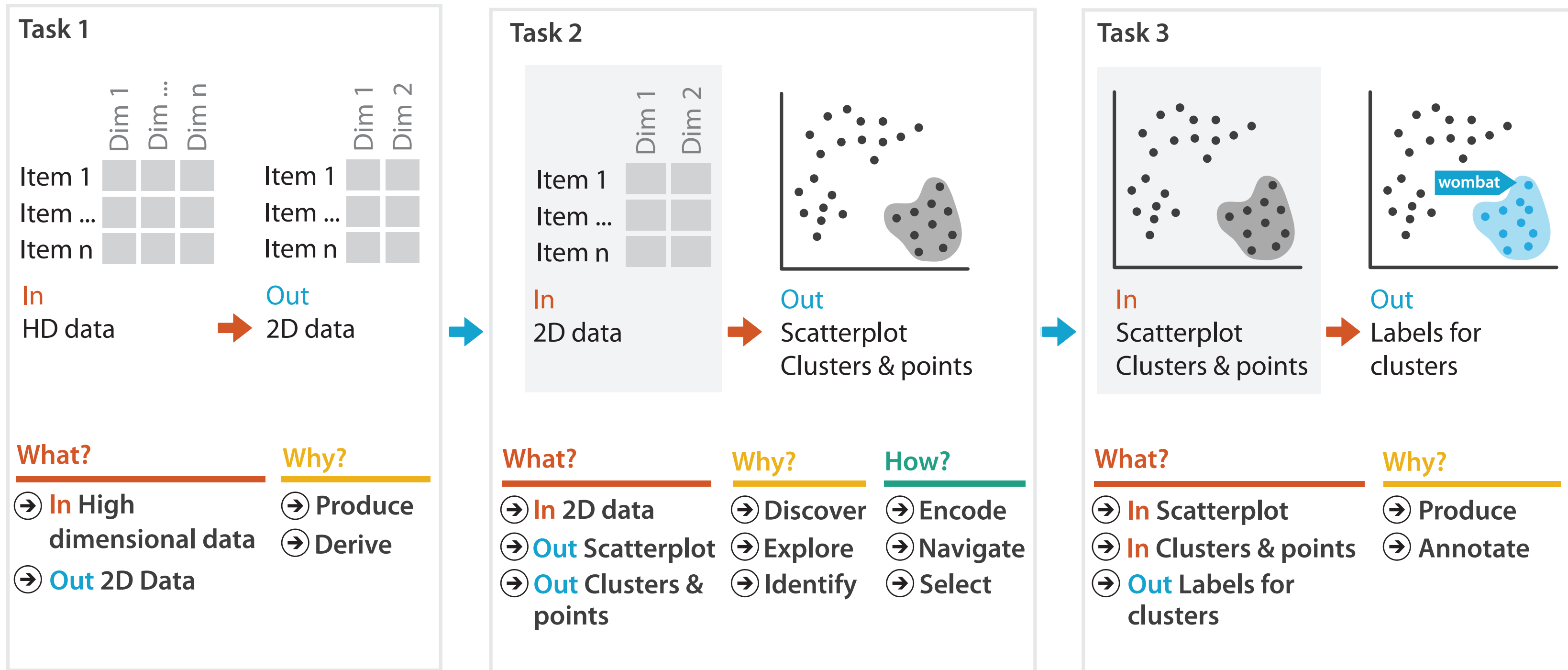


- Glimmer: multilevel dimensionality reduction algorithm
 - scalability to 30K documents and terms

*[Glimmer: Multilevel MDS on the GPU.
Ingram, Munzner, Olano. IEEE TVCG 15(2):249-261, 2009.]*



Starting point: Dimensionality reduction for document datasets



- more on DR: hour-long talk *Dimensionality Reduction from Several Angles*

<http://www.cs.ubc.ca/~tmm/talks.html#kelowna16>

Overview: Early version

The screenshot displays the MoDiscoTag software interface, which is divided into several functional panels:

- Disconnected Component Tree:** A hierarchical tree diagram on the left side, with a vertical axis labeled "Distance Threshold" ranging from 0 to 1.0. Below the tree is a "Show Nodes >=" control with buttons for 1, 2, 4, 8, 16, 32, and 64. The current view shows 107 nodes, with a sample list: "107: surinam dutch bouters guyana paramaribo jagdeo khan".
- Tags View:** A central panel with a "NEW" button and a list of "Selected Nodes" with associated controls. The list includes:

Oil industry	+	-	X
Finance	+	-	X
extradition	+	-	X
Banks	+	-	X
Airlines	+	-	X

A "SAVE" button is located at the bottom of this panel.
- Items Plot:** A scatter plot on the right showing a dense cluster of points, with some points highlighted in yellow and orange. Below the plot are two sliders: "Squeeze" and "Point Size", and a "Run..." button.
- Active Set List:** A bottom panel containing two columns of text, each preceded by a folder icon. The text represents active sets of nodes, such as "16: uav turkei shipment blue_lantern iran turkey\' ankara turk unscr item colr" and "10: uav turkei shipment iran turkey\' ankara unscr got related_materiel mater". A "CLEAR" button is positioned to the right of this list.

Overview: current version

OVERVIEW

Blog Help Contact us

admin@overview-project.org Admin ▾ Your document sets Log out

Search all documents Search ▾

+

-

service_offshore

investigation, independence, inc, atlantis_platform, production, whether, let MOST: l

ALL: investigation, atlantis_pl

MOST: hub, facility,

ALL: leases, royalty

ALL: investiga

ALL: c

ALL

ALL: indej

ALL

Tags atlantis form letter rig visit tag name Create new tag

organize tags...

Back to list

Document 1 of 11

Previous in folder MOST: hub, facility, canyon SOME: independence... Next


MMS2 Pdf 68 168 168

rig visit

Key words: arrive, port_fourchon, aviation_service, depart, shell Show sidebar

DOCUMENT PAGES TEXT Zoom Search

p. 1



Offshore Schedule

Friday, September 25, 2009

11:30 a.m.	Arrive Atlantic Aviation Service, 749 Lockheed Dr, Kenner, LA Gen. Mgr. Keith Myer, (Phone 504-453-8207); Receive helicopter safety briefing, personal protective equipment
12:00 noon	Depart Atlantic Aviation Service
1:15 p.m.	Arrive at Shell Offshore Inc.'s (Shell) Mars or Ursa Tension Leg Platform (TLP) located at Mississippi Canyon Block 807 or 809. Personnel introductions and personnel receive safety briefing/presentation by Shell.
2:00 p.m.	Tour either the Mars or Ursa TLP
4:00 p.m.	Depart TLP and flyover Louisiana Offshore Oil Port (LOOP) located at Grand Isle Block 59 enroute to Port Fourchon
5:00 p.m.	Arrive at Port Fourchon, Louisiana

Overview evolution: rationale?

The screenshot displays the Overview Project web application interface. At the top, a navigation bar includes the site name "OVERVIEW" and links for "Blog", "Help", and "Contact us". On the right, it shows the user's email "admin@overview-project.org", a dropdown menu for "Admin", and links for "Your document sets" and "Log out".

The main content area is divided into three sections:

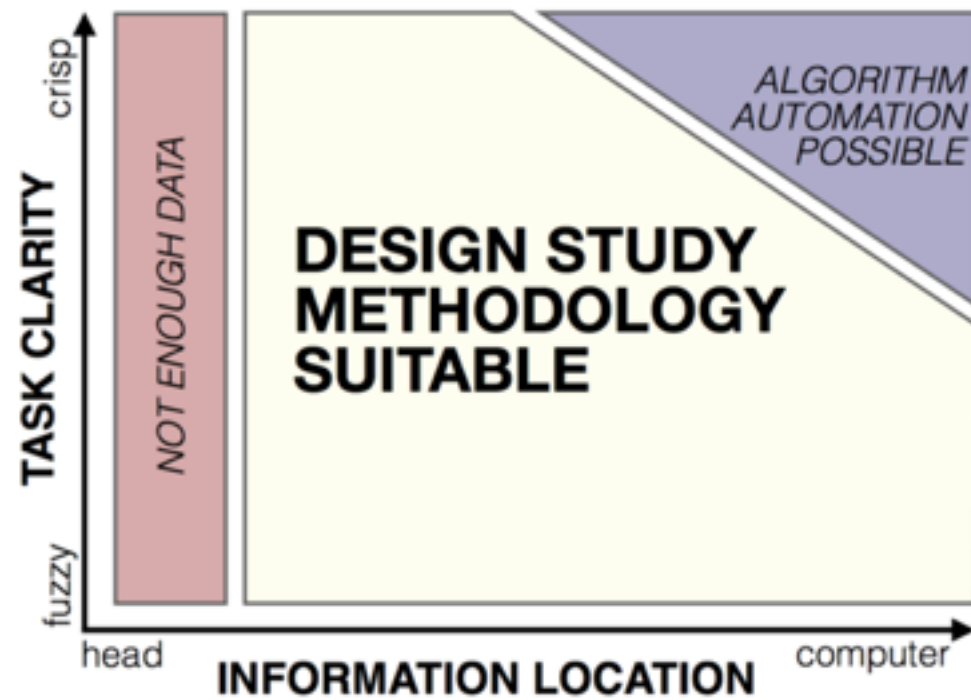
- Search:** A search bar at the top left contains the text "brad, letter, cheri, blythe, gas, oil, program, ocs, hunter, leases". Below it, a list of search results is shown, with the first result containing the text "letter, blythe_brad, decision, comments, director, june, july, contact".
- Folder tree:** A hierarchical tree structure is displayed. The root node is labeled "Folder tree" and contains sub-nodes such as "MOST: blythe_brad SOME: atlan...", "investigation, independence", "MOST: bly", and "MOST: br". The "investigation, independence" node is expanded to show further sub-nodes: "ALL: inves", "MOST", and "ALL: le".
- Document Viewer:** A document viewer on the right side shows "Document 9 of 40" in a folder named "investigation, independence, inc, atlantis_platform, pr...". The document title is "MMS1 Pdf 24 130 131" with a tag "atlantis". Key words listed include "investigation, atlantis_platform, monicasandersmailhousegov, anita_atlantic included_mms", remembering". The document content is an email from Brad J. Blythe to Monica Sanders, dated Wednesday, July 29, 2009. The email text includes: "My apologies for not remembering this while we were on the phone. Due to the nature of this request, I have included MMS' statement on inquiries into the Atlantis Platform. The Minerals Management Service received a copy of a letter from a special interest group to the Department of the Interior requesting an investigation into British Petroleum's Atlantis platform, which is operating in the Gulf of Mexico. MMS is currently reviewing the contents of the letter. As a matter of policy, however, we do not publicly discuss whether investigations are ongoing or pending in order to maintain the integrity of the investigation process. I realize this probably isn't very helpful at the moment, but this is all we are allowed to say for now. -Brad". The sender's name and title are "Brad J. Blythe, Ph.D. Presidential Management Fellow" and the organization is "Department of the Interior Minerals Management Service Offshore Energy and Minerals Management".

At the bottom of the interface, there is a "Tags" section with a list of tags: "atlantis", "contains", "environment", "impact", "form letter", and "rig visit". A "Create new tag" button and a "tag name" input field are also visible.

Deploy in the real world



Case Study	#1	#2	#3	#4	#5	#6
Document Collection	4,500 pages from FOIA	5,996 emails from FOIA	8,680 pages from FOIA	1,278 survey comments	4,653 emails from FOIA	1,680 bills
Question	<i>What did security contractors do during Iraq war?</i>	<i>Were municipal police funds mismanaged?</i>	<i>Were Paul Ryan's campaign statements hypocritical?</i>	<i>What is the gun ownership debate about?</i>	<i>Was gov't response to emergency incident effective?</i>	<i>Did gov't fail to pass bills addressing police misconduct?</i>



Michael Sedlmair



Miriah Meyer

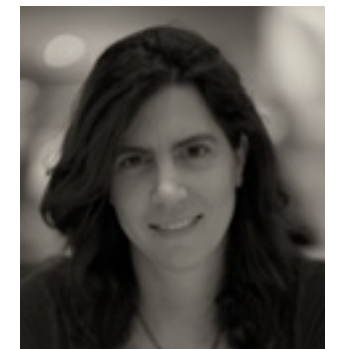


Design Study Methodology

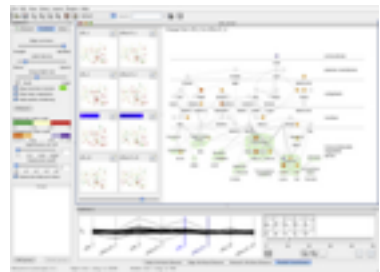
Reflections from the Trenches and from the Stacks

<http://www.cs.ubc.ca/labs/imager/tr/2012/dsm/>

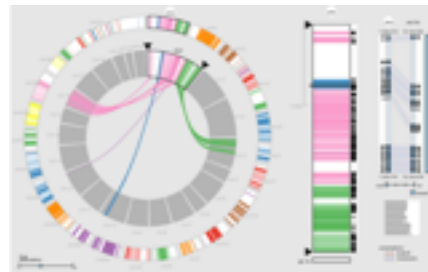
Tamara Munzner
@tamaramunzner



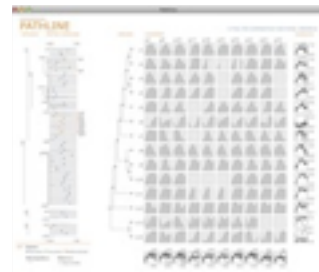
Design Studies: Lessons learned after 21 of them



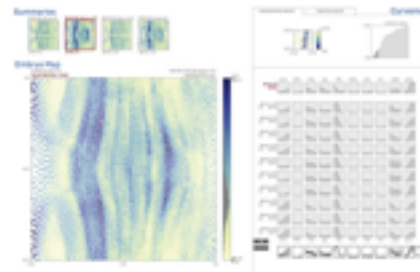
Cerebral
genomics



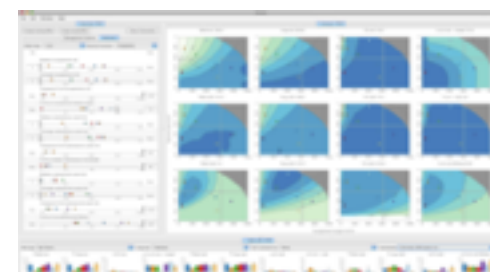
MizBee
genomics



Pathline
genomics



MulteeSum
genomics



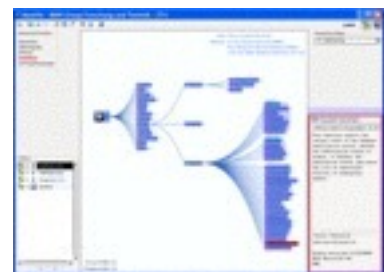
Vismon
fisheries management



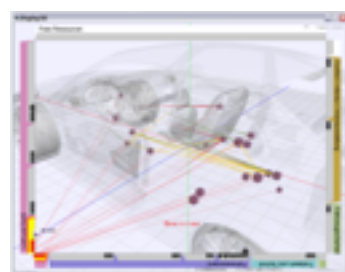
QuestVis
sustainability



WiKeVis
in-car networks



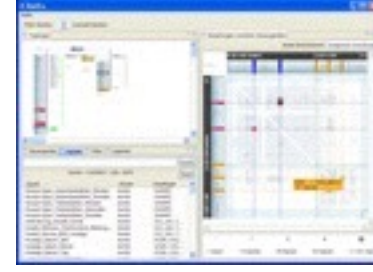
MostVis
in-car networks



Car-X-Ray
in-car networks



ProgSpy2010
in-car networks



ReEx
in-car networks



Cardiogram
in-car networks



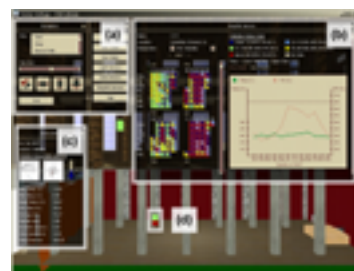
AutobahnVis
in-car networks



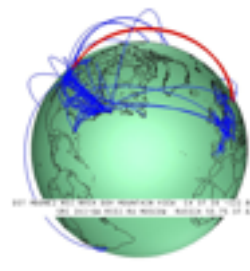
VisTra
in-car networks



Constellation
linguistics



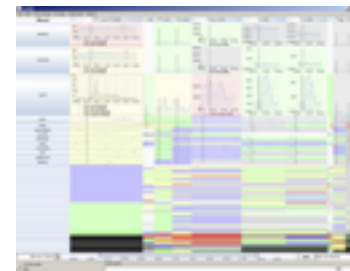
LibVis
cultural heritage



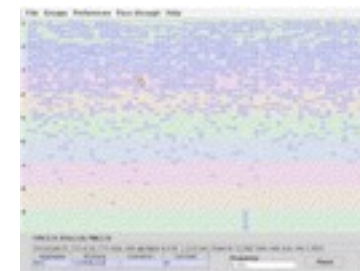
Caidants
multicast



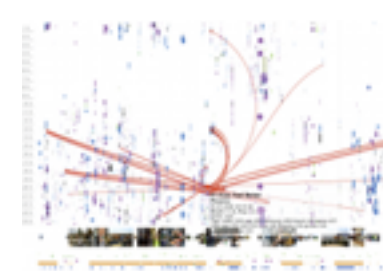
SessionViewer
web log analysis



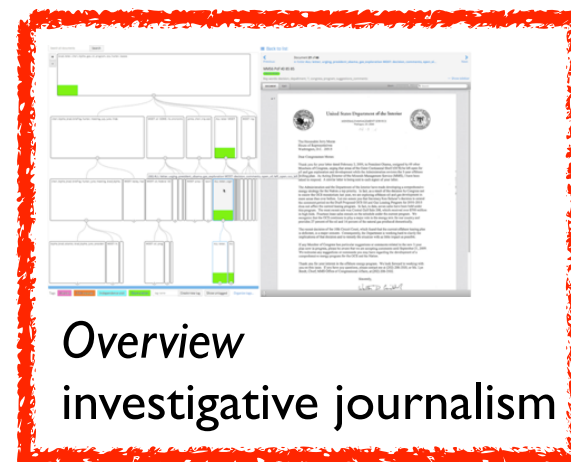
LiveRAC
server hosting



PowerSetViewer
data mining

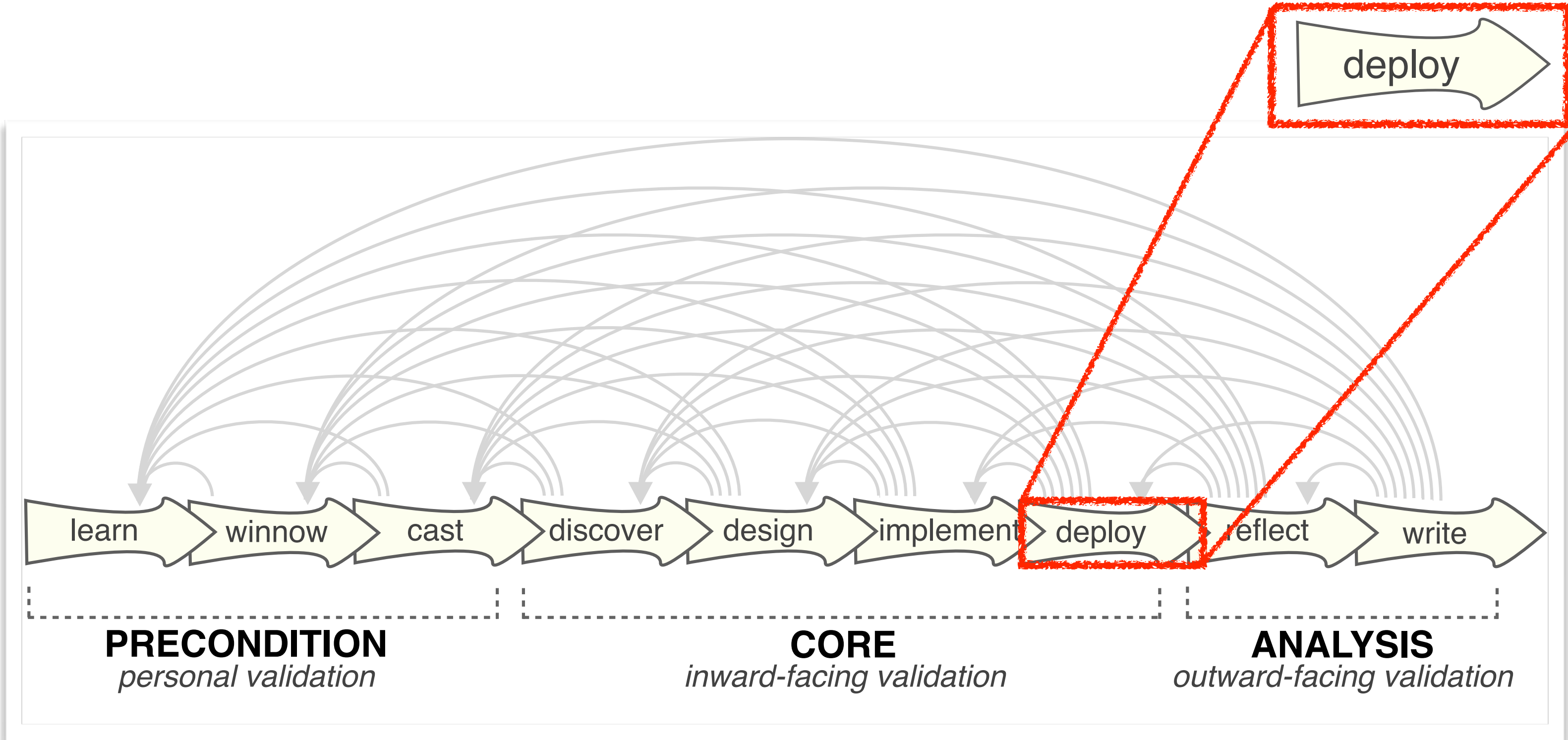


LastHistory
music listening

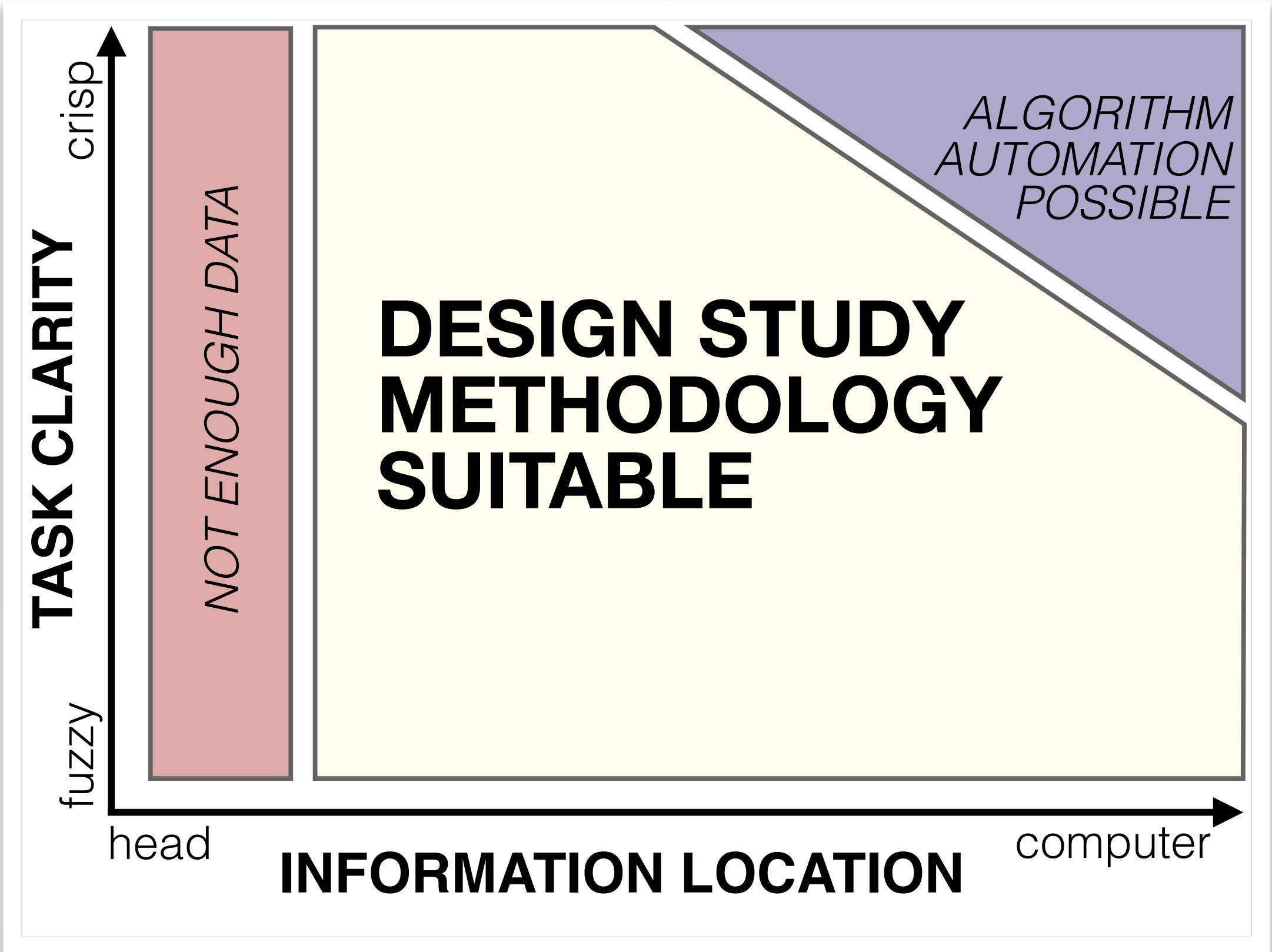


Overview
investigative journalism

Design study methodology: 9-stage framework



Design study methodology: definitions



Design study methodology: 32 Pitfalls

- and how to avoid them

PF-1	premature advance: jumping forward over stages	general
PF-2	premature start: insufficient knowledge of vis literature	learn
PF-3	premature commitment: collaboration with wrong people	winnow
PF-4	no real data available (yet)	winnow
PF-5	insufficient time available from potential collaborators	winnow
PF-6	no need for visualization: problem can be automated	winnow
PF-7	researcher expertise does not match domain problem	winnow
PF-8	no need for research: engineering vs. research project	winnow
PF-9	no need for change: existing tools are good enough	winnow

Collaboration incentives

- why do CS/vis people need to understand journalism's problems?
 - we work with you to understand your driving problems
 - we build tools intended to help
 - only works out if we understood the problems deeply enough
 - we observe how you use them
 - if they're good enough
 - CS win: research success stories
 - journalist win: access to better tools
 - we develop guidelines on how to build better tools in general
 - CS win: research progress in visualization

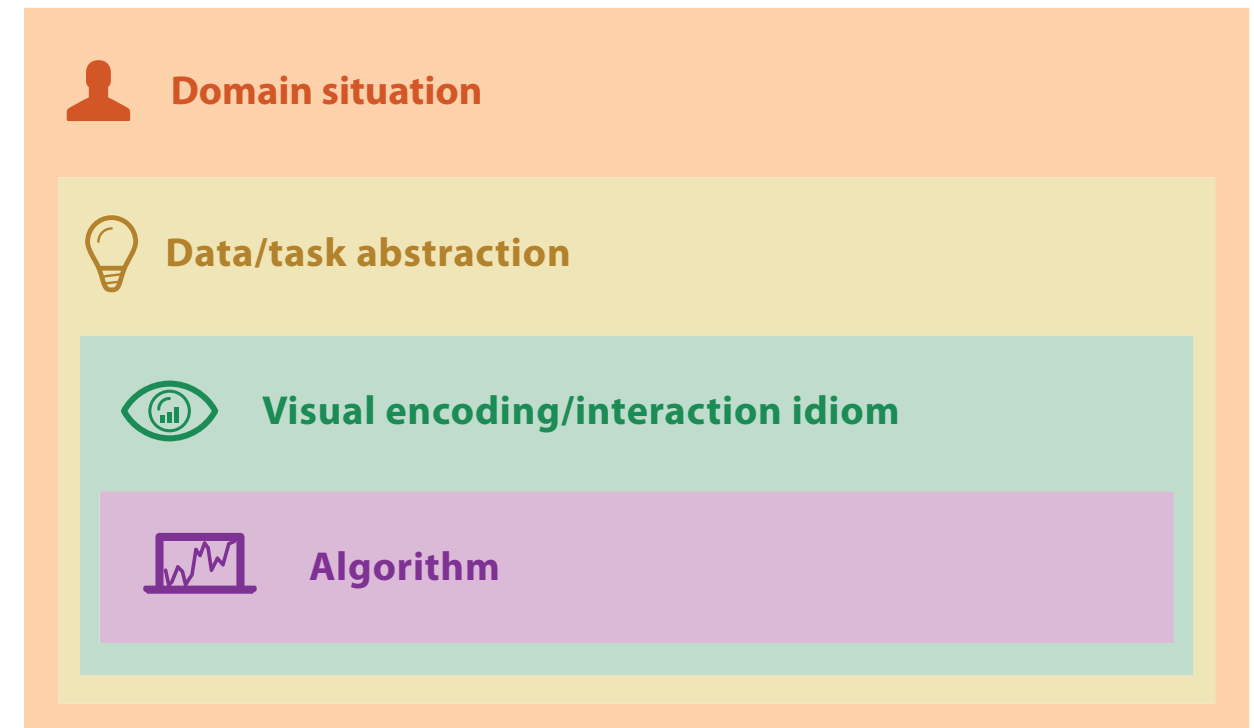
Deploy in the real world, understand user goals

Case Study	#1	#2	#3	#4	#5	#6
Document Collection	4,500 pages from FOIA	5,996 emails from FOIA	8,680 pages from FOIA	1,278 survey comments	4,653 emails from FOIA	1,680 bills
Question	<i>What did security contractors do during Iraq war?</i>	<i>Were municipal police funds mismanaged?</i> find the needle in the haystack	<i>Were Paul Ryan's campaign statements hypocritical?</i>	<i>What is the gun ownership debate about?</i>	<i>Was gov't response to emergency incident effective?</i>	<i>Did gov't fail to pass bills addressing police misconduct?</i> prove haystack contains no needles!

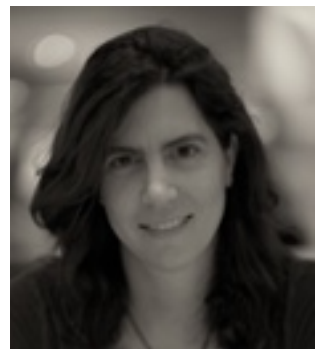
A Nested Model

for Visualization Design and Validation

www.cs.ubc.ca/labs/imager/tr/2009/NestedModel



Tamara Munzner
@tamaramunzner

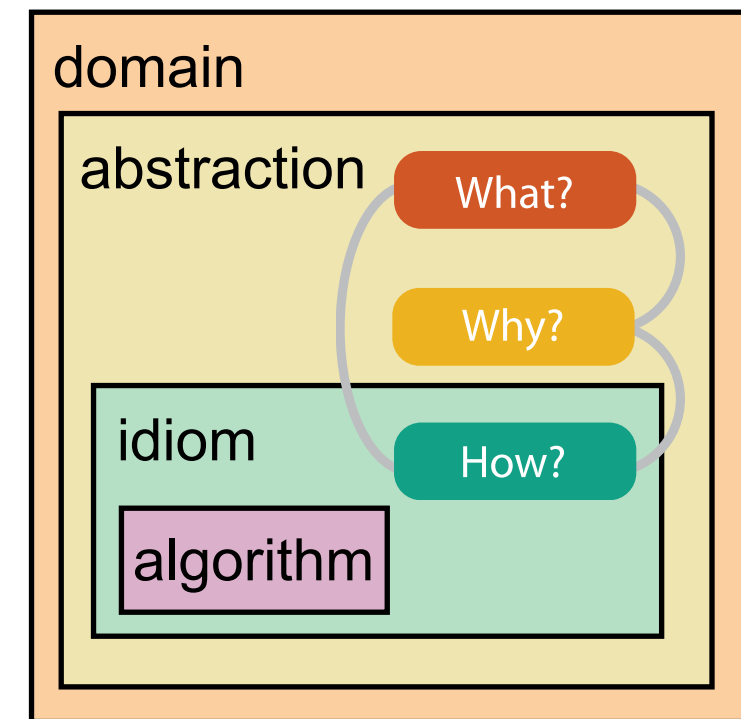


A Nested Model for Visualization Design and Validation.
Munzner. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09)*, 15(6):921-928, 2009.

Nested model: Four levels of vis design

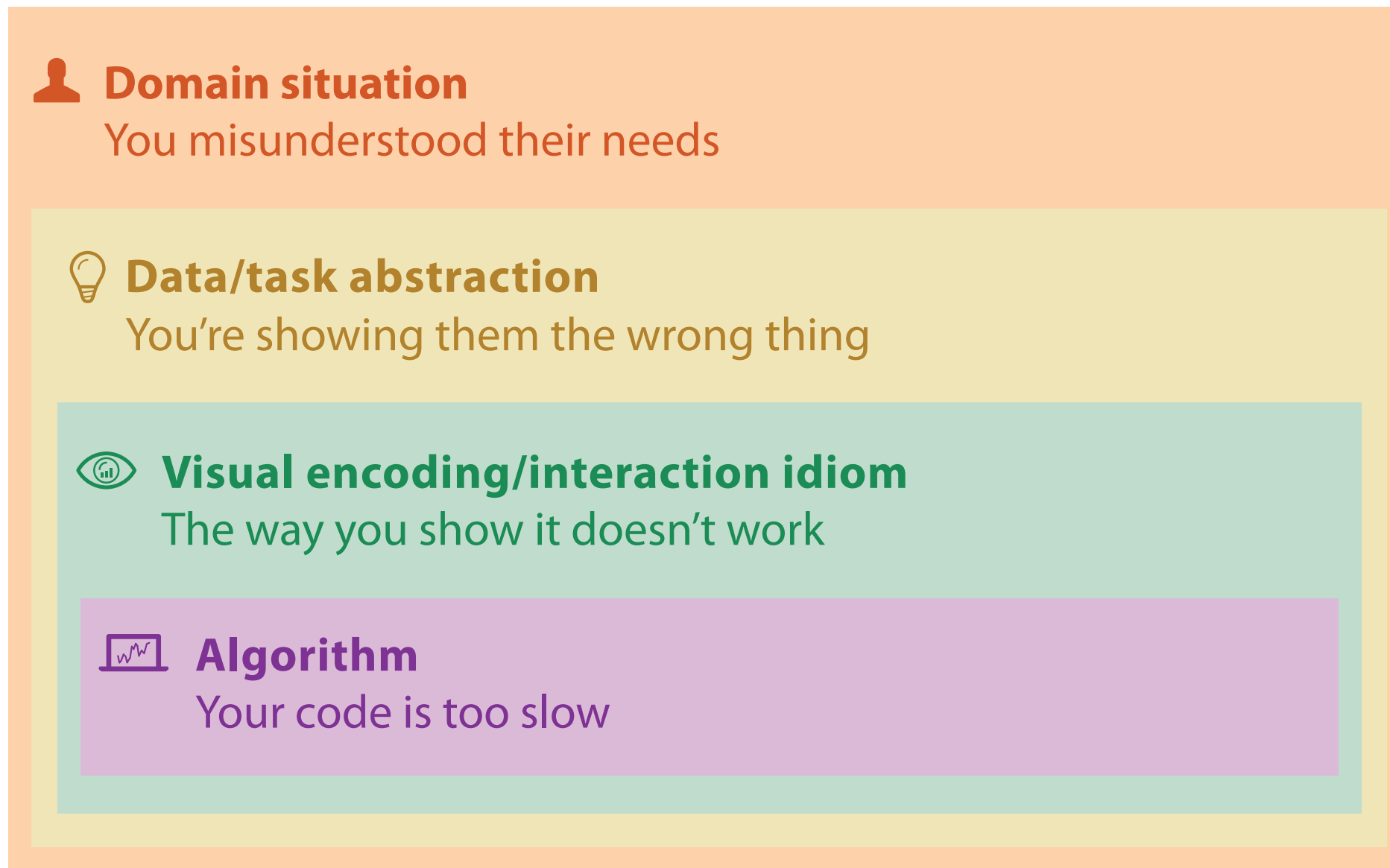
- *domain situation*
 - who are the target users?
 - CS: domain = journalism; journ: domain = story topic
- *abstraction*
 - translate from specifics of domain to vocabulary of vis
 - **what** is shown? data abstraction
 - **why** is the user looking at it? task abstraction
- *idiom*
 - **how** is it shown?
 - **visual encoding** idiom: how to draw
 - **interaction** idiom: how to manipulate
- *algorithm*
 - efficient computation

[A Nested Model of Visualization Design and Validation.
Munzner. *IEEE TVCG* 15(6):921-928, 2009
(Proc. InfoVis 2009).]

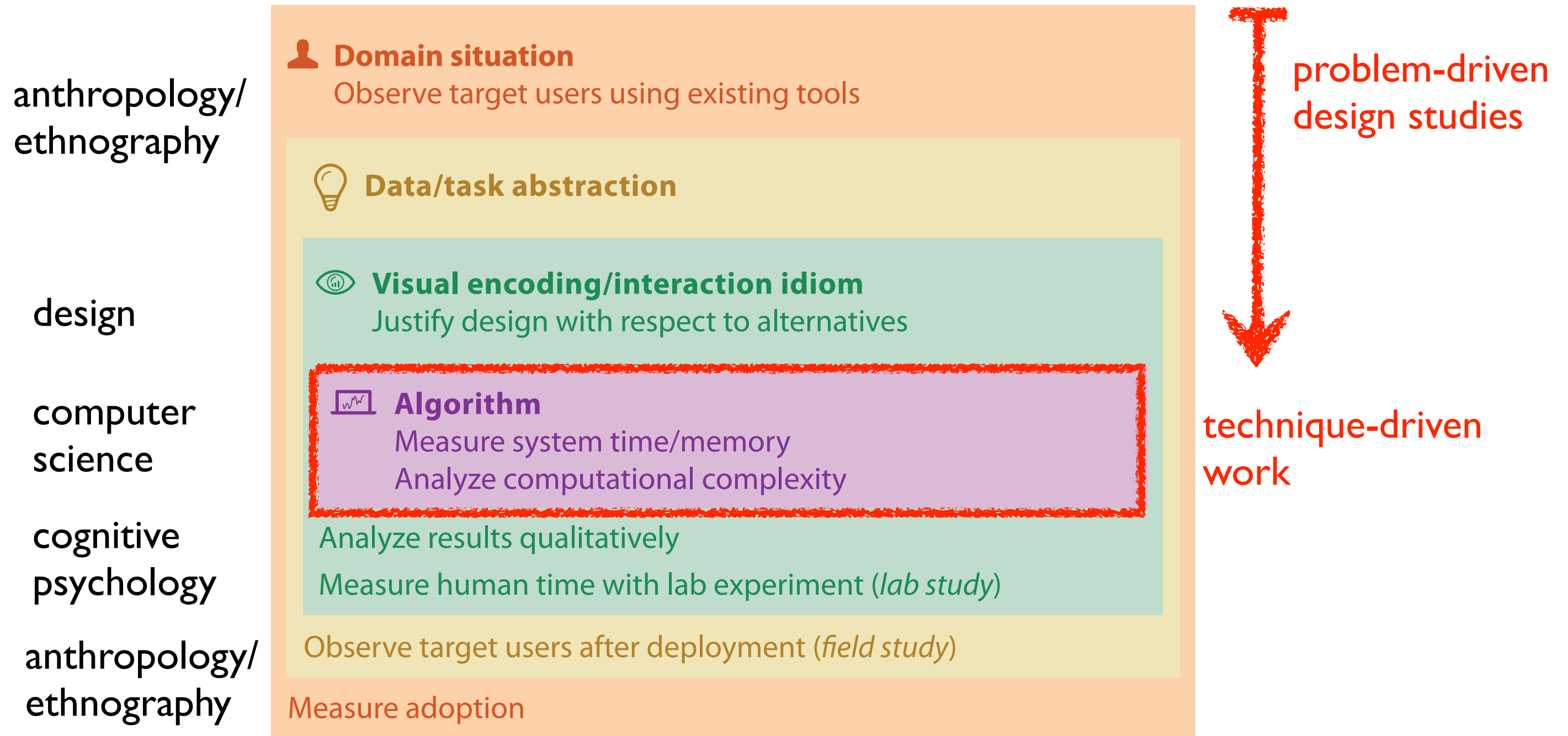


[A Multi-Level Typology of Abstract Visualization Tasks
Brehmer and Munzner. *IEEE TVCG* 19(12):2376-2385,
2013 (Proc. InfoVis 2013).]

Threats to validity differ at each level



Evaluate success at each level with methods from different fields



Evolution across levels

- evolution of task abstraction

- task 1: **generate hypotheses** → **explore** → **summarize**

- *obviously you can't read everything; speed up with tool for categorizing and counting*

- task 2: **verify hypotheses** → **locate** → **identify**

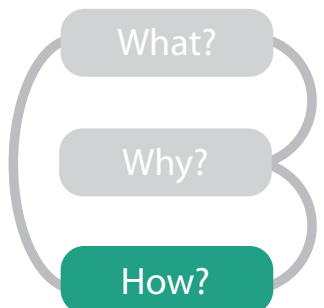
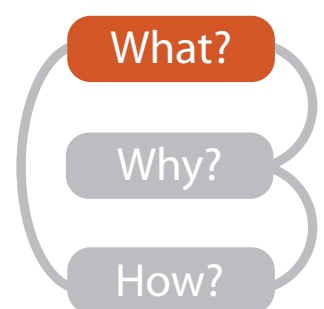
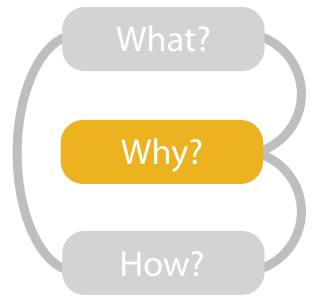
- *you really do read each doc; speed up with tool to keep track of findings*

- evolution of data abstraction & idioms

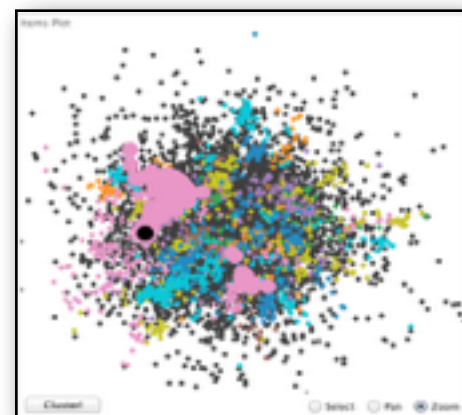
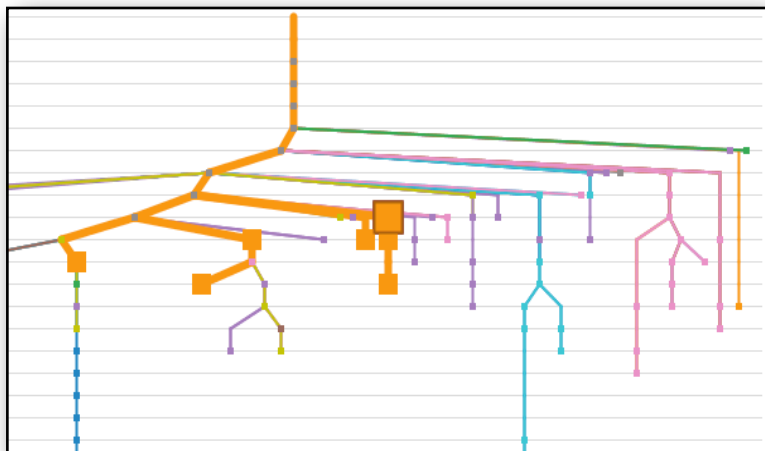
- arrange cluster tree to emphasize nodes vs links

- new vis insight: DR scatterplot less effective than cluster tree vis + tagging

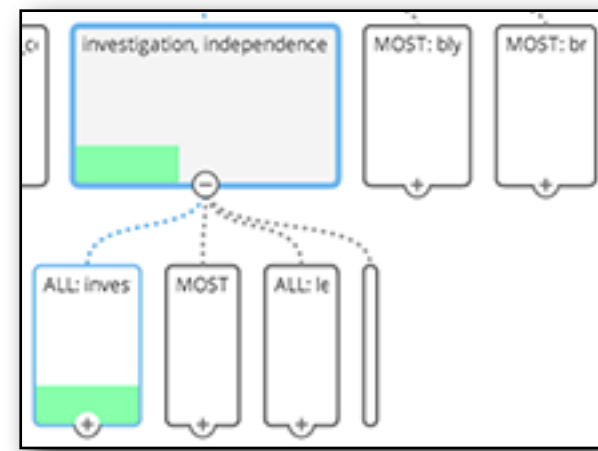
- better affordance for systematic traversal of document cluster hierarchy



early

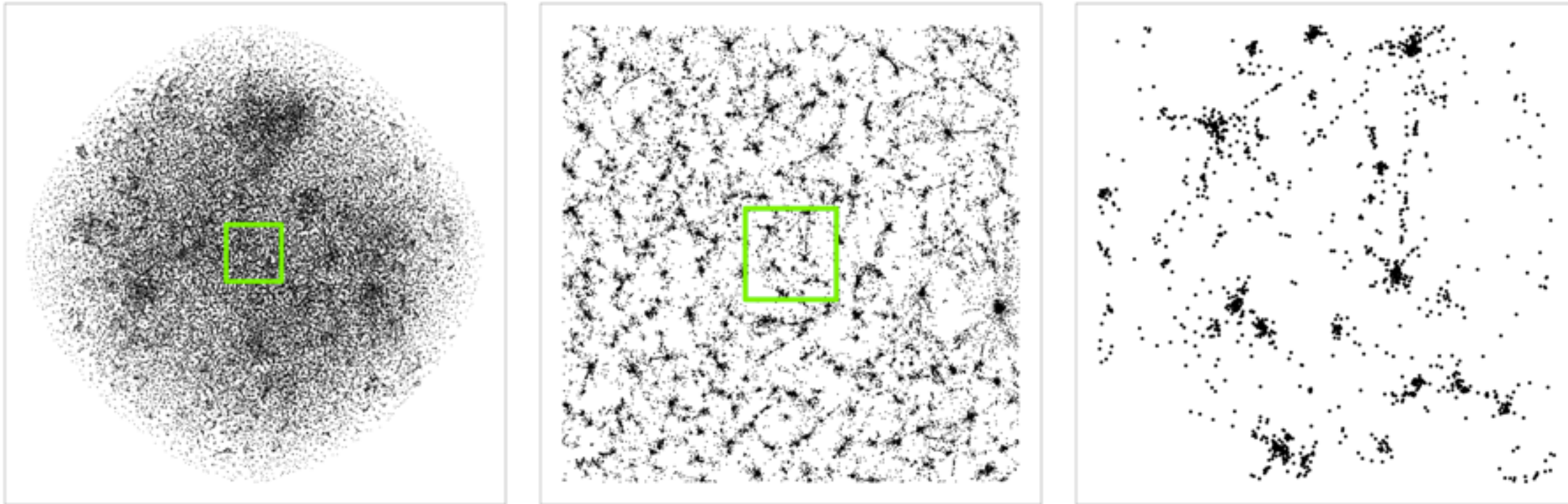
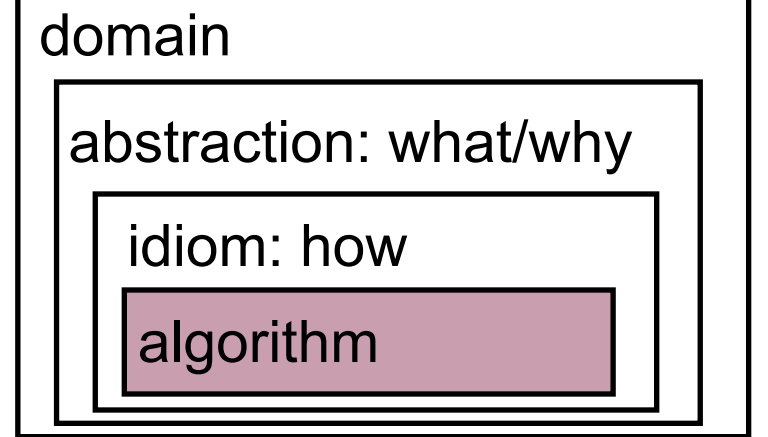


current



Algorithm: Spinoff series

- dimensionality reduction for huge text collections
 - great algorithm problem in its own right!
 - QSNE: fast and high-quality DR for millions of documents
 - key feature: handle sparseness appropriately



[Dimensionality Reduction for Documents with Nearest Neighbor Queries. Ingram and Munzner. Neurocomputing (Special Issue on Visual Analytics using Multidimensional Projections), Volume 150 Part B, p 557-569, 2015.]

<http://www.cs.ubc.ca/labs/imager/tr/2014/QSNE/>

Vignette 2: Vis Tool for Journalistic Presentation



Johanna Fulda
@jofu_



Matthew Brehmer
@mattbrehmer



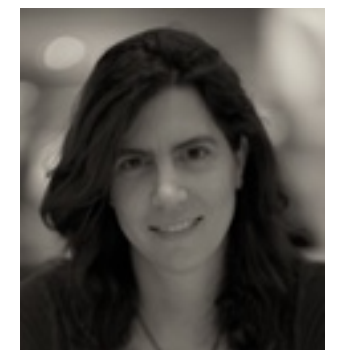
TimeLineCurator

Interactive Authoring of Visual Timelines from Unstructured Text

<http://about.timelinecurator.org>

<http://timelinecurator.org>

Tamara Munzner
@tamaramunzner



TimeLineCurator: Interactive Authoring of Visual Timelines from Unstructured Text.
Fulda, Brehmer, Munzner. *IEEE Trans. Visualization and Computer Graphics (Proc IEEE VAST 2015)* 22(1):300-309, 2015.

Origin story: Tedium in the newsroom

- Johanna Fulda: interactive infographics developer, Sueddeutsche Zeitung
 - then Munich CS master's student, visiting UBC
- what pain point could we address with interactive visualization?
 - plus some NLP
- sound familiar?...

TimeLineCurator

visual & browser-based

<https://vimeo.com/jofu/tlc>

Manual creation process



1868 The Typewriter
 Invented by Christopher Sholes, typewriters quickly became indispensable tools for practically all writing other than personal correspondence. They were widely used by professional writers, in offices, and for business correspondence in private homes.

1868
 Christopher Sholes
 Typewriter

1977
 Douglas C. Engelbart
 Mouse
 + One Button Computer

1986
 The Mouse
 Some additional information here

1997
 The Stylus
 a small pen-shaped instrument that is used to input commands to a computer screen, mobile device or graphics tablet.

2007
 Multi Touch
 With the start of iPhones Multi-touch became a thing

2012
 Speech Recognition

Mighty Mouse
 In 1980, Apple Computer asked a group of guys fresh from Stanford's product design program to take a \$400 device and make it mass-producible, reliable and cheap. Their work transformed personal computing.















Structured creation process



	A	B	C	D	E	F
	Start Date	End Date	Headline	Text	Media	Media Credit
1				Zuckerberg wrote a program called Facemash on October 28, 2003 while attending Harvard as a sophomore	http://dubindigital.ie/wp-content/uploads/2011/03/facemash-2003-2009.jpg	Dublin Digital
2	10/28/2003		Facemash	The following semester, Zuckerberg began writing code for a new website in January 2004		
3	1/1/2004		new website	On February 4, 2004, Zuckerberg launched "thefacebook", originally located at thefacebook.com		
4	2/4/2004		thefacebook.com	Six days after the site launched, three Harvard seniors (Cameron Winklevoss, Tyler Winklevoss, and Divya Narendra) accused Zuckerberg of intentionally misleading them into believing he would help them build a social network called HarvardConnection.com	http://www.capitoberg.com/wp-content/uploads/2011/05/3Harvard-wider.jpg	Capital Berg
5	2/10/2004		Harvard Connection accusations	They later filed a lawsuit against Zuckerberg, subsequently settling in 2008[17] for 1.2 million shares (worth \$300 million at Facebook's IPO)	http://4.bp.blogspot.com/_KMKs4_C0jys/YY5336_YFsi/AAAAAAAAAT8s/bitao-c5Q3TAx26005.jpeg	Estynese
6	1/1/2008	12/31/2008	Harvard Connection	Membership was initially restricted to students of Harvard College; within the first month, more than half the undergraduates at Harvard were registered on the service		
7	2/4/2004	3/4/2004	thefacebook at Harvard only	In March 2004, Facebook expanded to the universities of Columbia, Stanford, and Yale [20]		
8	3/1/2004	3/31/2004	thefacebook expands to other universities	In mid-2004, entrepreneur Sean Parker (an informal advisor to		

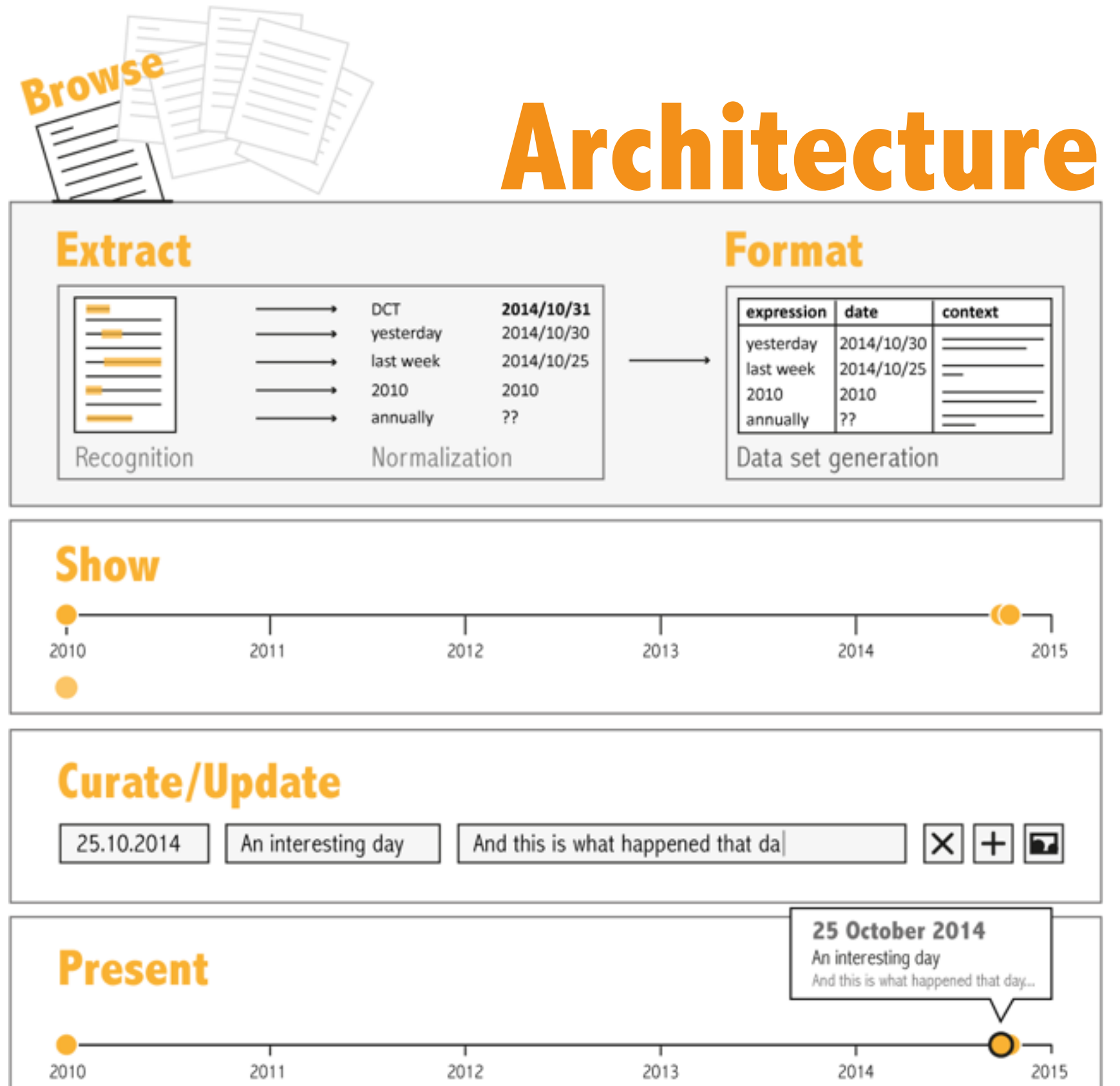
Timeline authoring model

- time required for each task

	Browse	Extract	Format	Show	Update
Manual Drawing	 slow	 slow		 slow	 slow
Structured Creation	 slow	 slow	 slow	 automated	 fast
TimeLine Curator	 fast	 automated	 automated	 fast	 fast

The general case for curation

- build for human in the loop as continuing need
 - automatic processing to accelerate not replace
 - assume computational results good but not perfect
 - for the indefinite future!
 - visual feedback to accelerate



The importance of being brisk

- sexy use case: eureka moment
 - enable what was impossible before
 - vis tools for new insights & discoveries
- workhorse use case: workflow speedup
 - vis tools to accelerate what you're already doing
 - sometimes enables the previously infeasible
- TLC use cases
 - started with speedup use case, for presentation
 - make this doc into a timeline now!
 - two other use cases nudge towards exploration
 - comparison between multiple timelines
 - speculative browsing



TimeLineCurator: Speculative Browsing

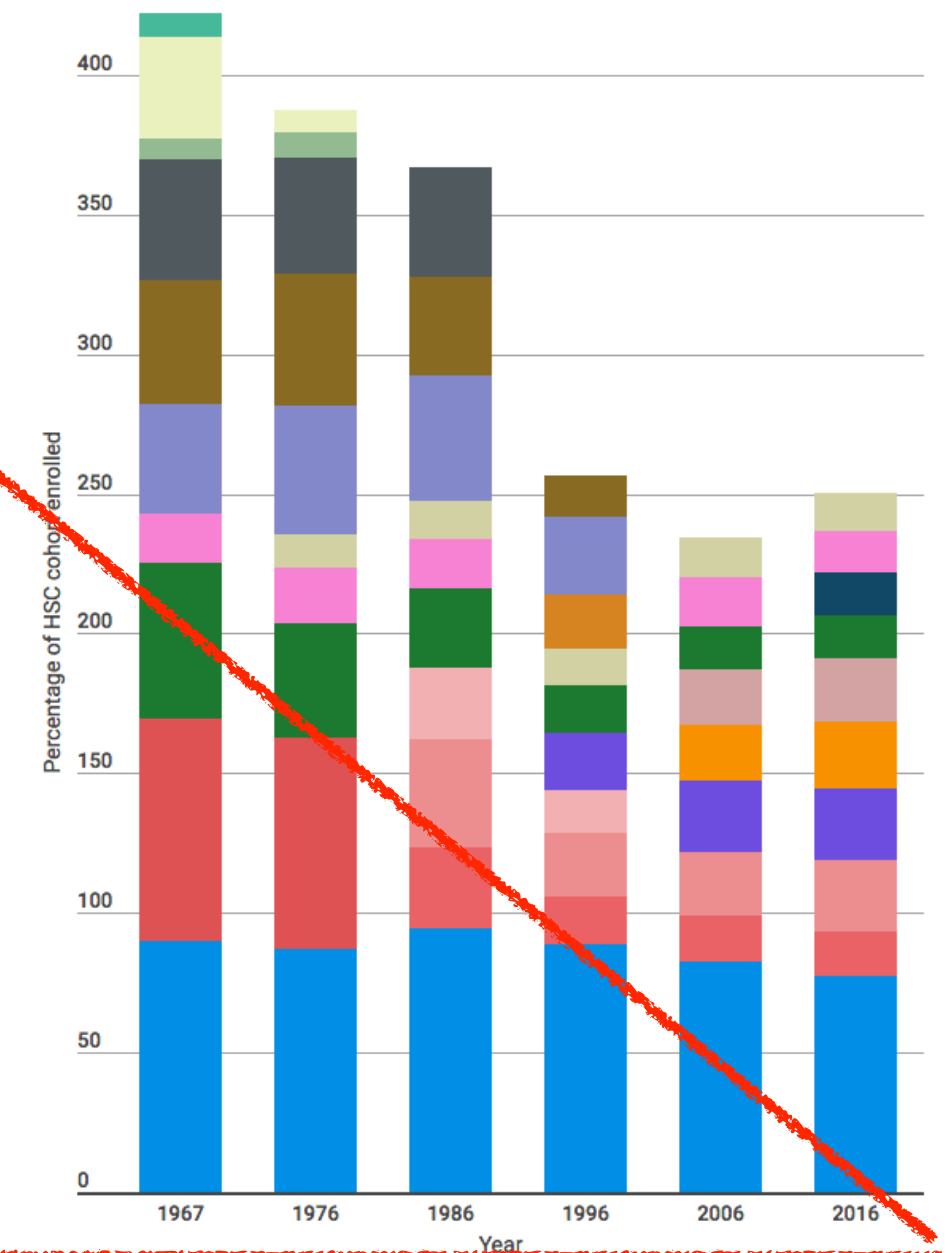
s p e c u l a t i v e b r o w s i n g

**Vignette 3:
Challenges of Color
(A Cautionary Tale)**

Challenges of Color

- what is wrong with this picture?

Top 10 HSC subjects (excluding English)



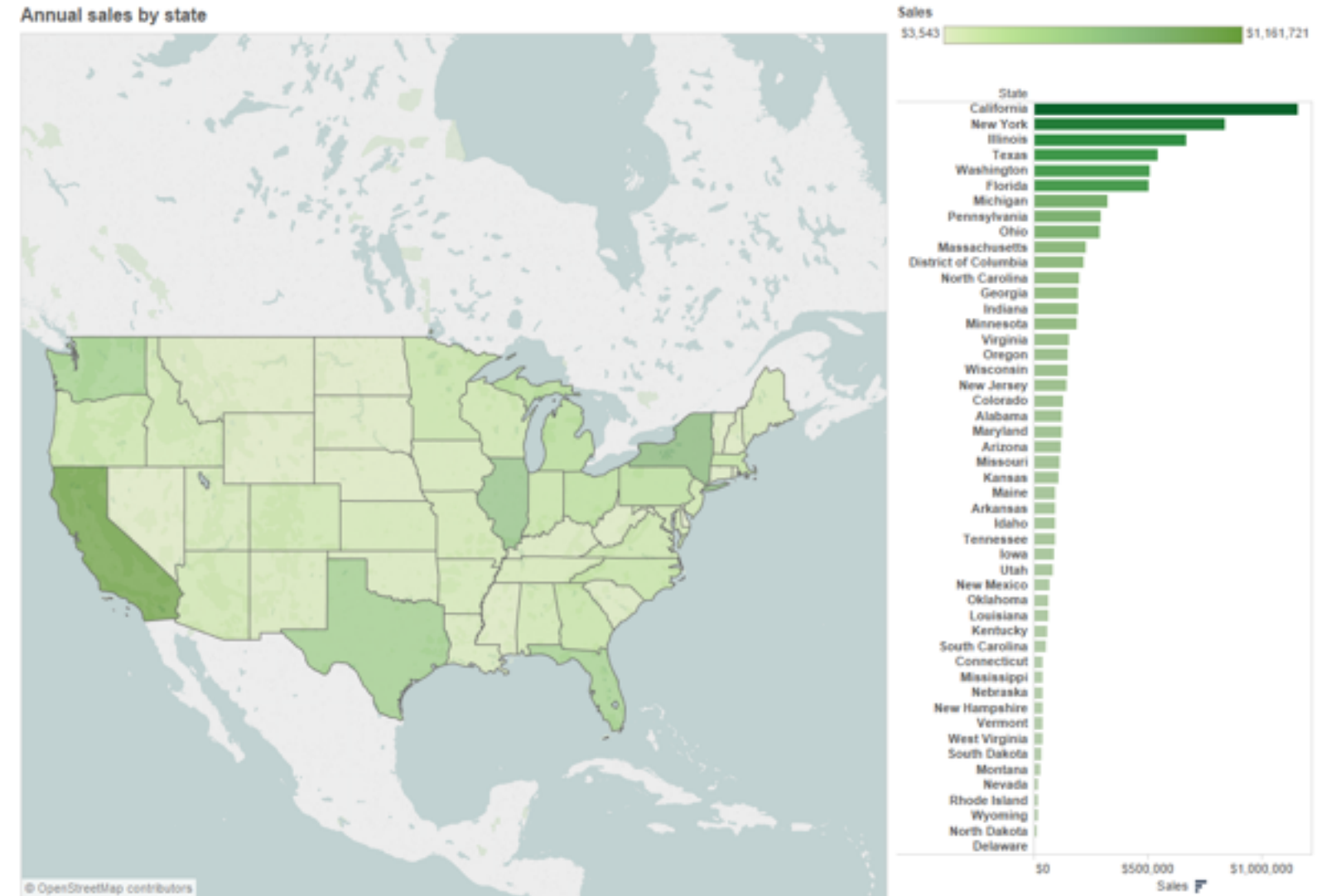
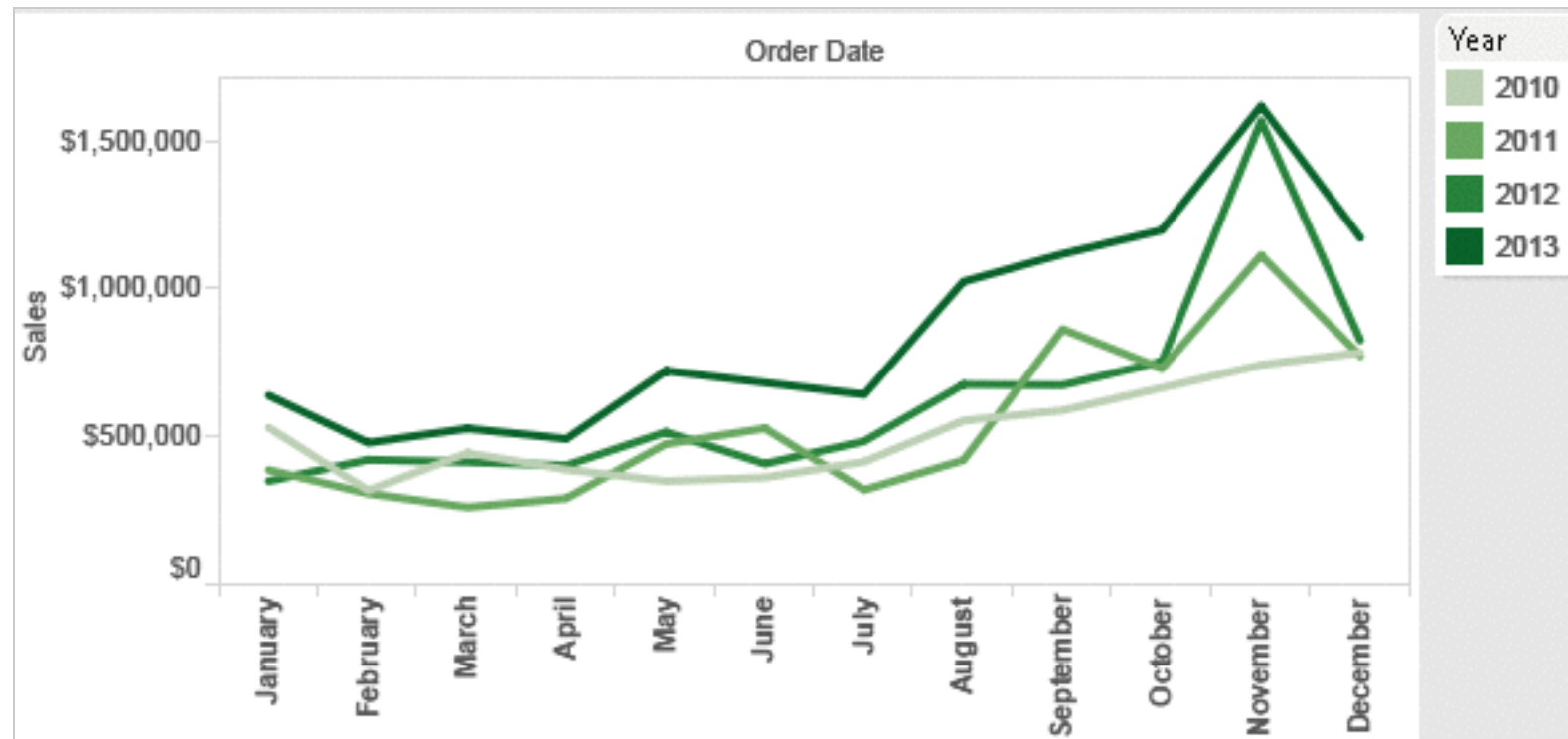
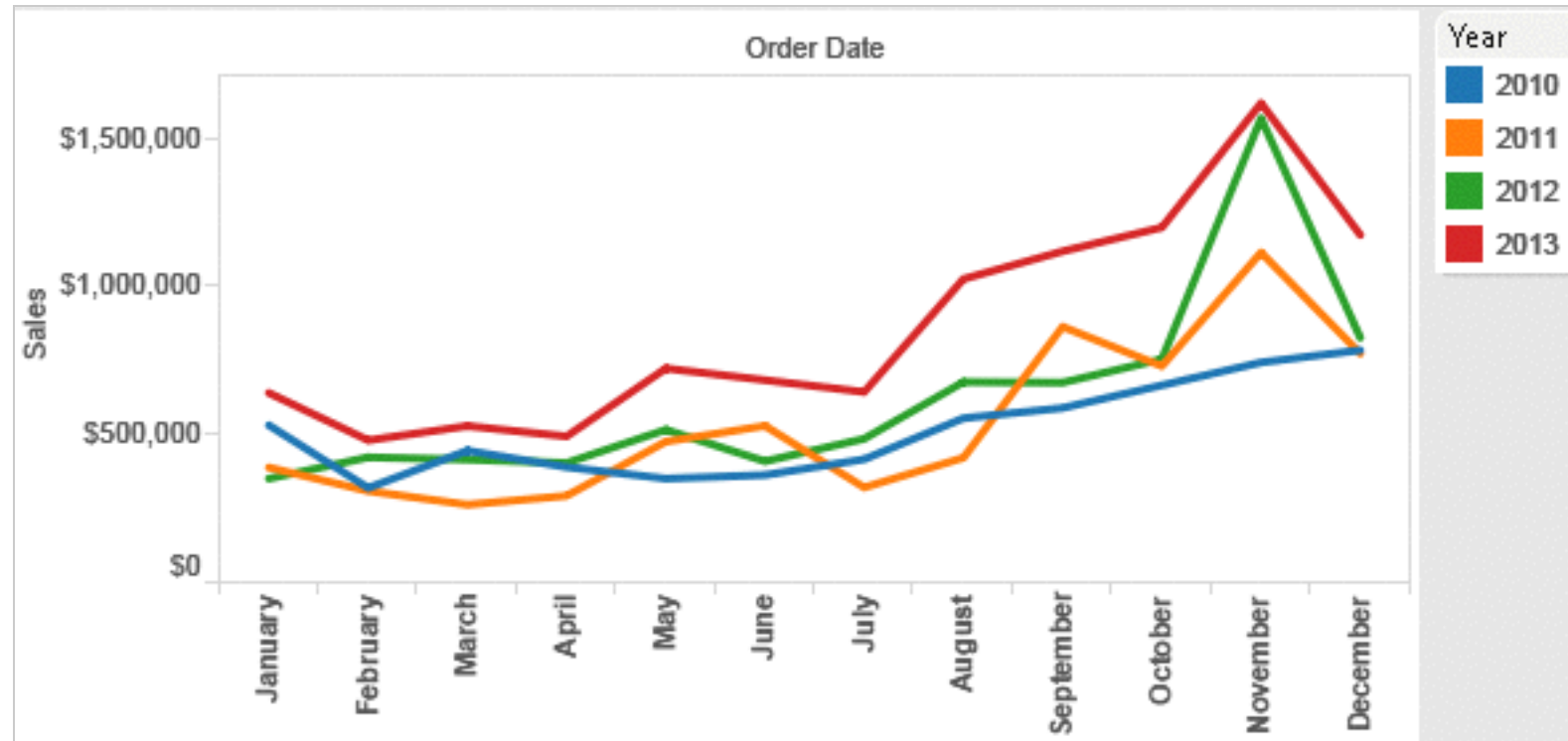
- Maths ● Science ● Chemistry ● Biology ● Physics ● Business Studies
- PDHPE ● Studies of Religion ● Modern History ● Legal Studies
- Ancient History ● Visual Arts ● Computing Studies ● General Studies
- Geography ● Economics ● Industrial Arts ● French ● Latin

@WTFViz

“visualizations that make no sense”

- Maths ● Science ● Chemistry ● Biology ● Physics ● Business Studies
- PDHPE ● Studies of Religion ● Modern History ● Legal Studies
- Ancient History ● Visual Arts ● Computing Studies ● General Studies
- Geography ● Economics ● Industrial Arts ● French ● Latin

Categorical vs ordered color



[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

Decomposing color

- first rule of color: do not talk about color!
 - color is confusing if treated as monolithic

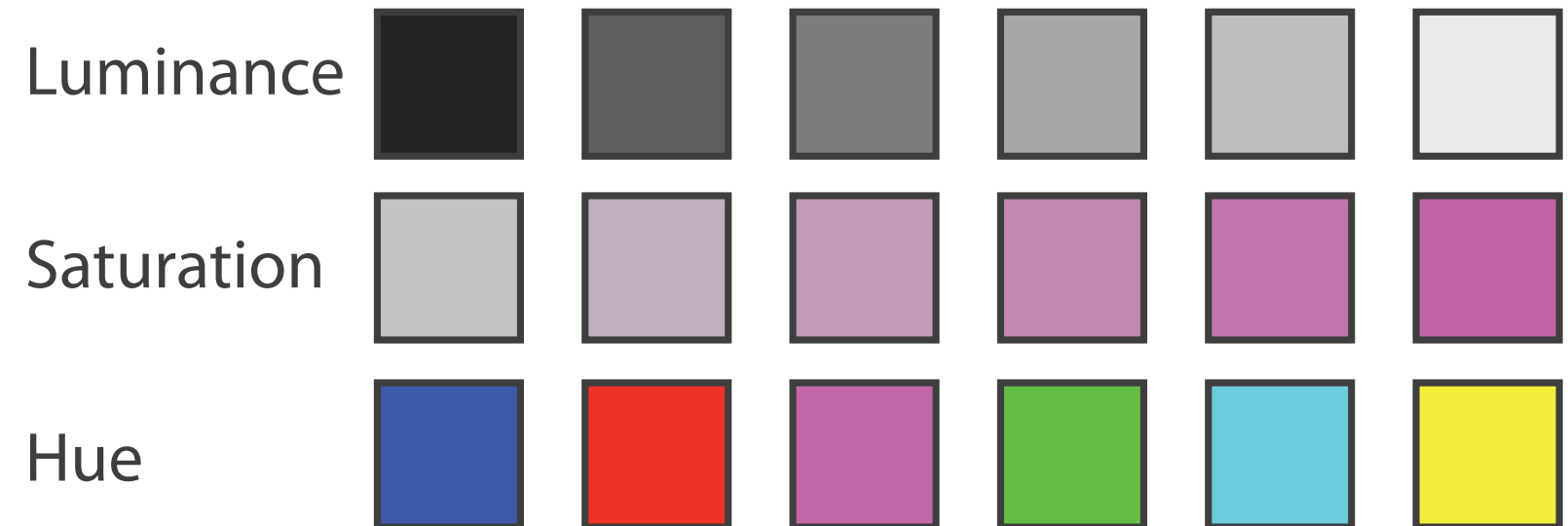
- decompose into three channels

- ordered can show magnitude

- luminance
- saturation

- categorical can show identity

- hue



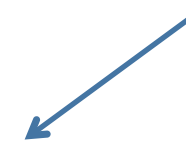
- channels have different properties

- what they convey directly to perceptual system

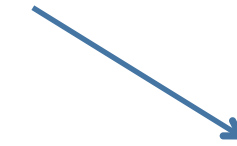
- how much they can convey: how many discriminable bins can we use?

Luminance

- need luminance for edge detection
 - fine-grained detail only visible through luminance contrast
 - legible text requires luminance contrast!
- intrinsic perceptual ordering



Lightness information



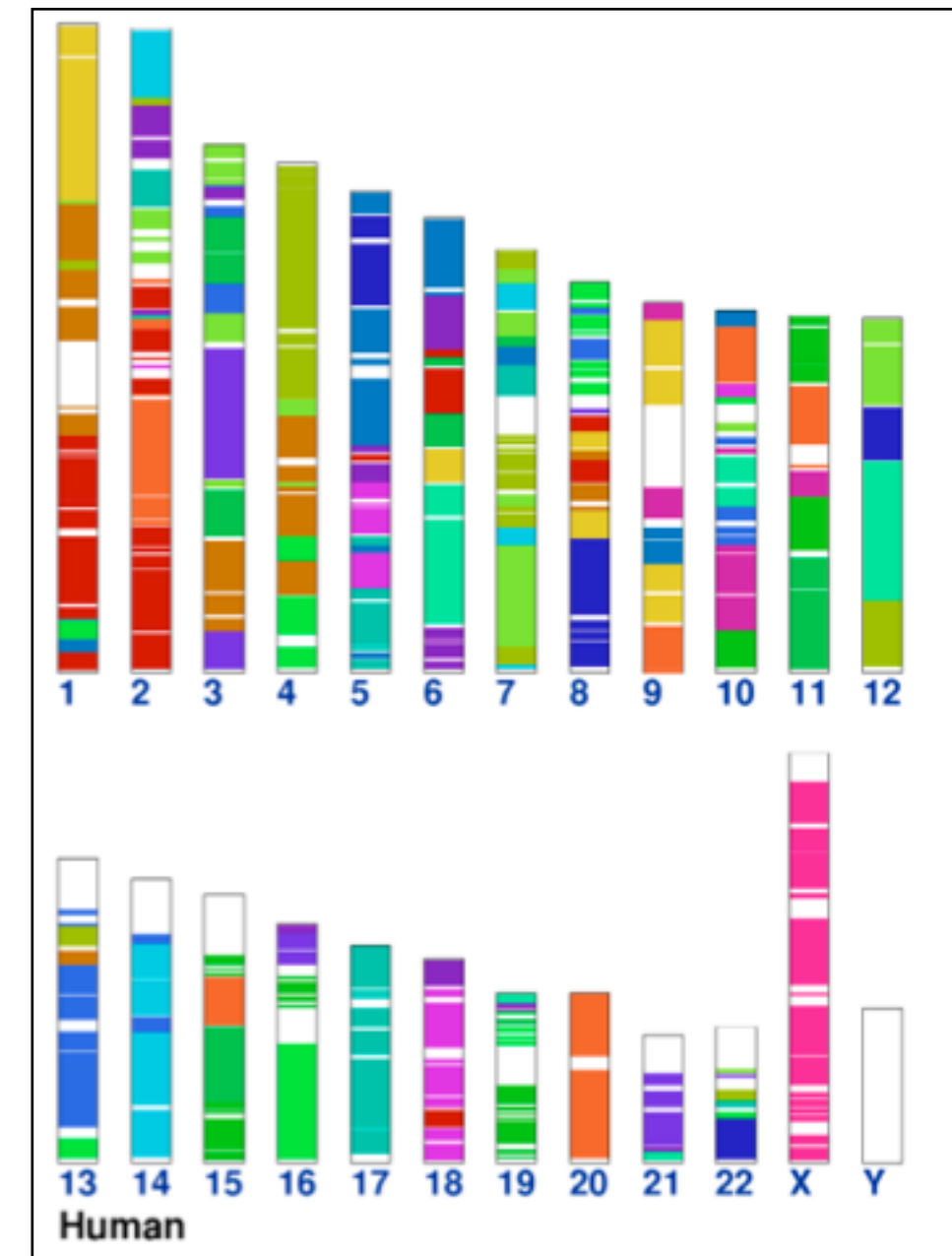
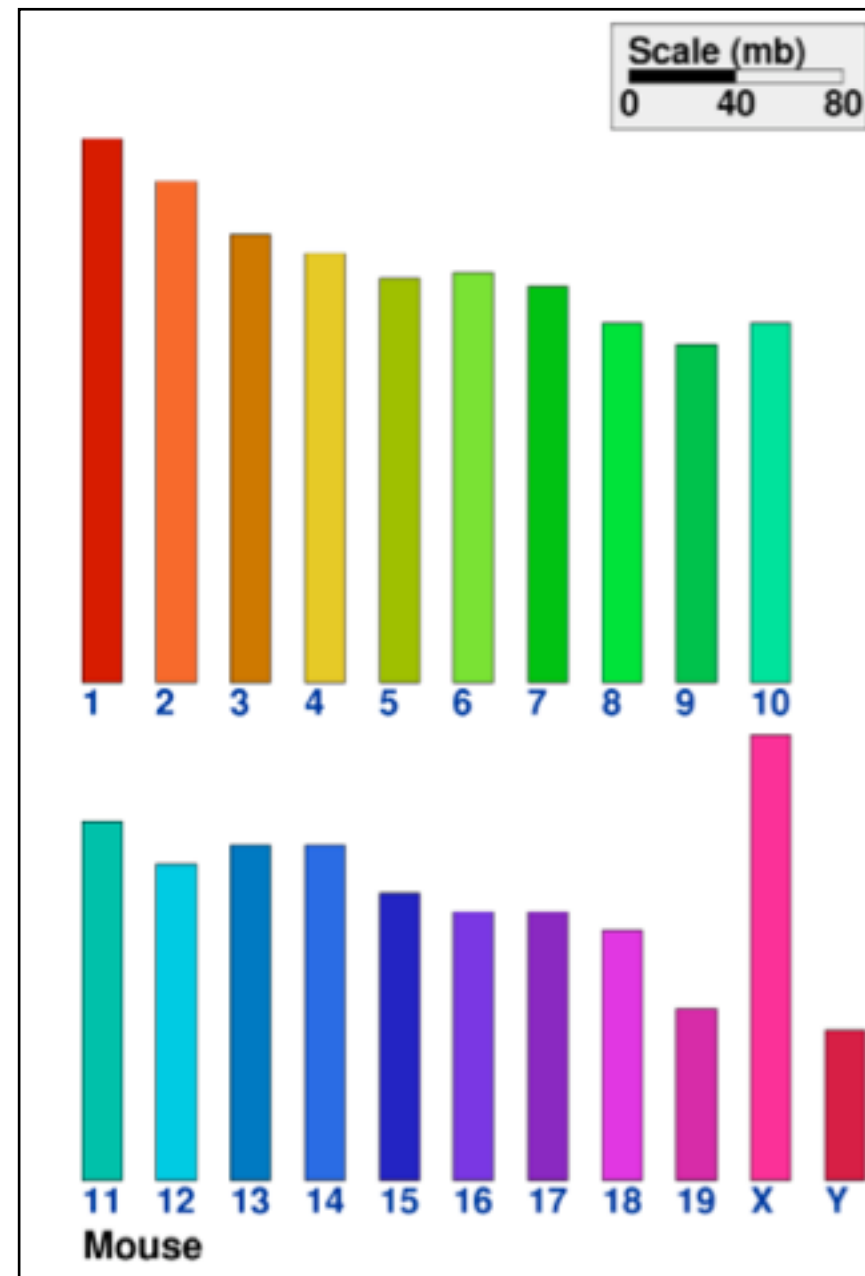
Color information



*[Seriously Colorful: Advanced Color Principles & Practices.
Stone.Tableau Customer Conference 2014.]*

Categorical color: limited number of discriminable bins

- human perception built on relative comparisons
 - great if color contiguous
 - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
 - fewer bins than you want
 - rule of thumb: 6-12 bins, including background and highlights



–so what can we do instead?

[Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

Analyzing visual encoding via marks and channels

- marks

- geometric primitives

- channels

- control appearance of marks

- channel properties differ

- type & amount of information that can be conveyed to human perceptual system

- number of discriminable bins

- show magnitude vs. identity

- accuracy of perception

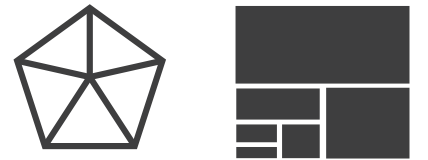
➔ Points



➔ Lines



➔ Areas



➔ Position

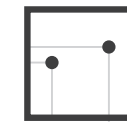
➔ Horizontal



➔ Vertical



➔ Both



➔ Color



➔ Shape



➔ Tilt



➔ Size

➔ Length



➔ Area





➔ Volume




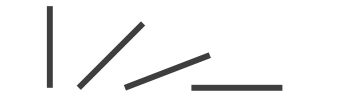
Channels: Matching expressiveness


➔ Magnitude Channels: Ordered Attributes


Position on common scale 


Position on unaligned scale 


Length (1D size) 


Tilt/angle 


Area (2D size) 

Depth (3D position) 

Color luminance 

Color saturation 

Curvature 

Volume (3D size) 

Same

Same

➔ Identity Channels: Categorical Attributes

Spatial region 

Color hue 

Motion 

Shape 

- expressiveness principle
 - match channel and data characteristics

➔ Attribute Types

➔ Categorical

➔ Ordered



➔ Ordinal

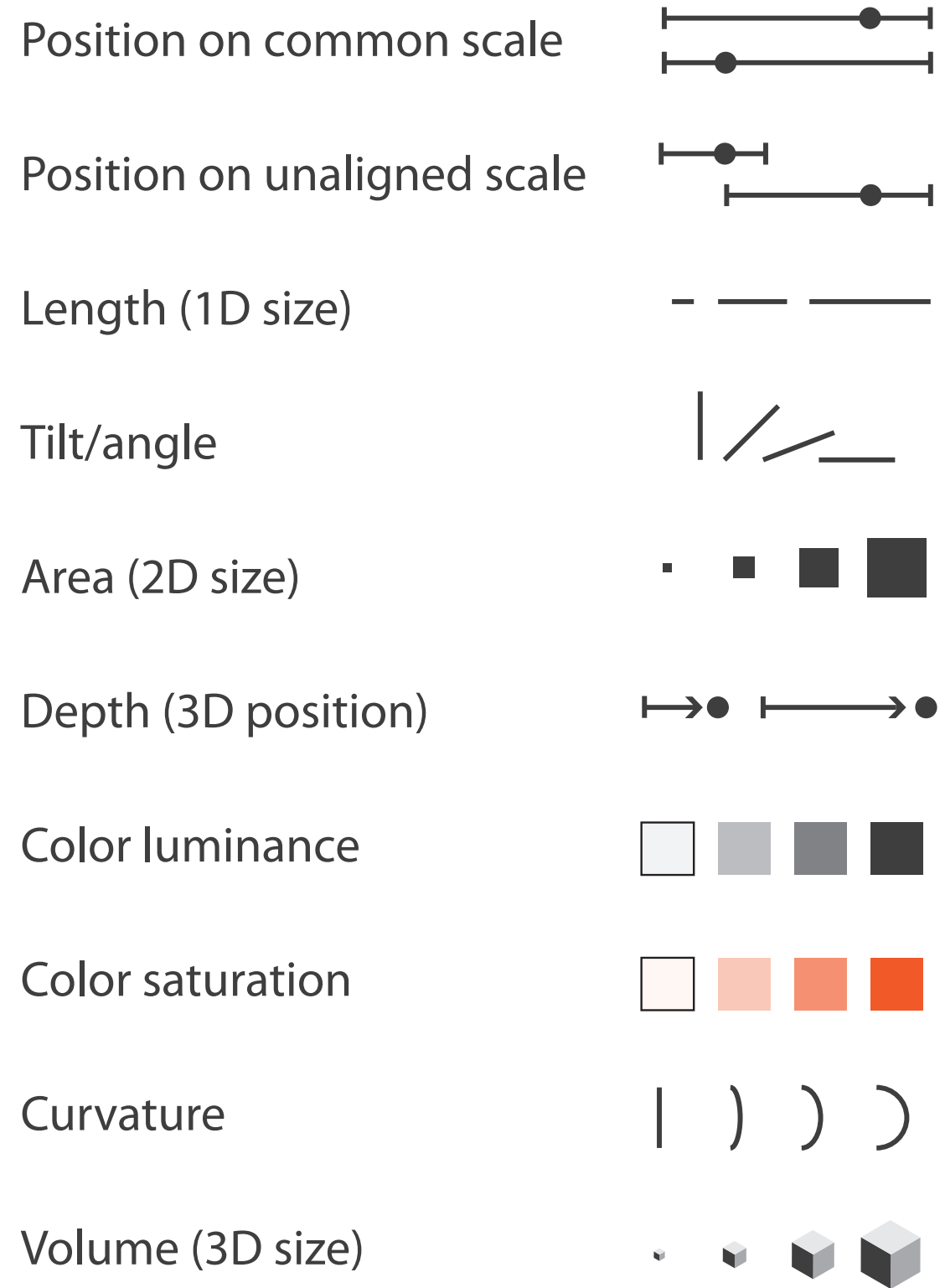


➔ Quantitative



Channels: Ranking effectiveness

➔ Magnitude Channels: Ordered Attributes



➔ Identity Channels: Categorical Attributes

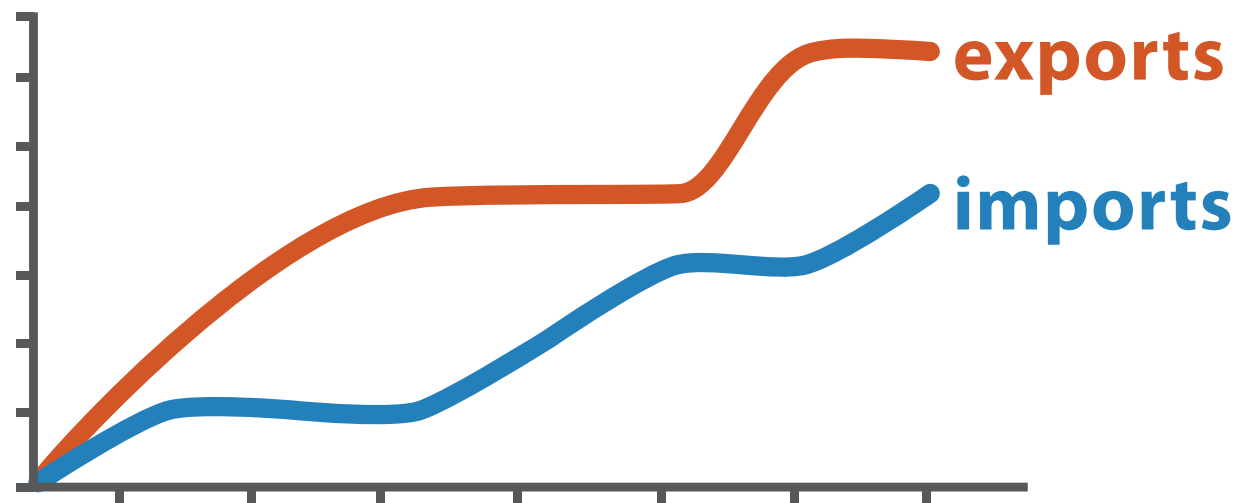


Best
Effectiveness
Least

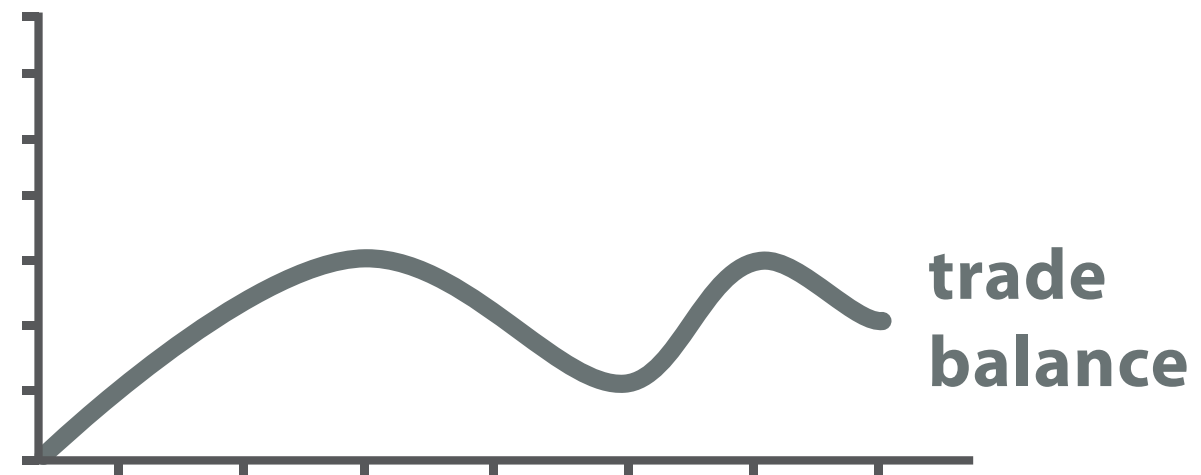
- **expressiveness principle**
 - match channel and data characteristics
- **effectiveness principle**
 - encode most important attributes with highest ranked channels

Derive

- don't just draw what you're given!
 - decide what the right thing to show is
 - create it with a series of transformations from the original dataset
 - draw that
- one of the four major strategies for handling complexity



Original Data



$$\text{trade balance} = \text{exports} - \text{imports}$$

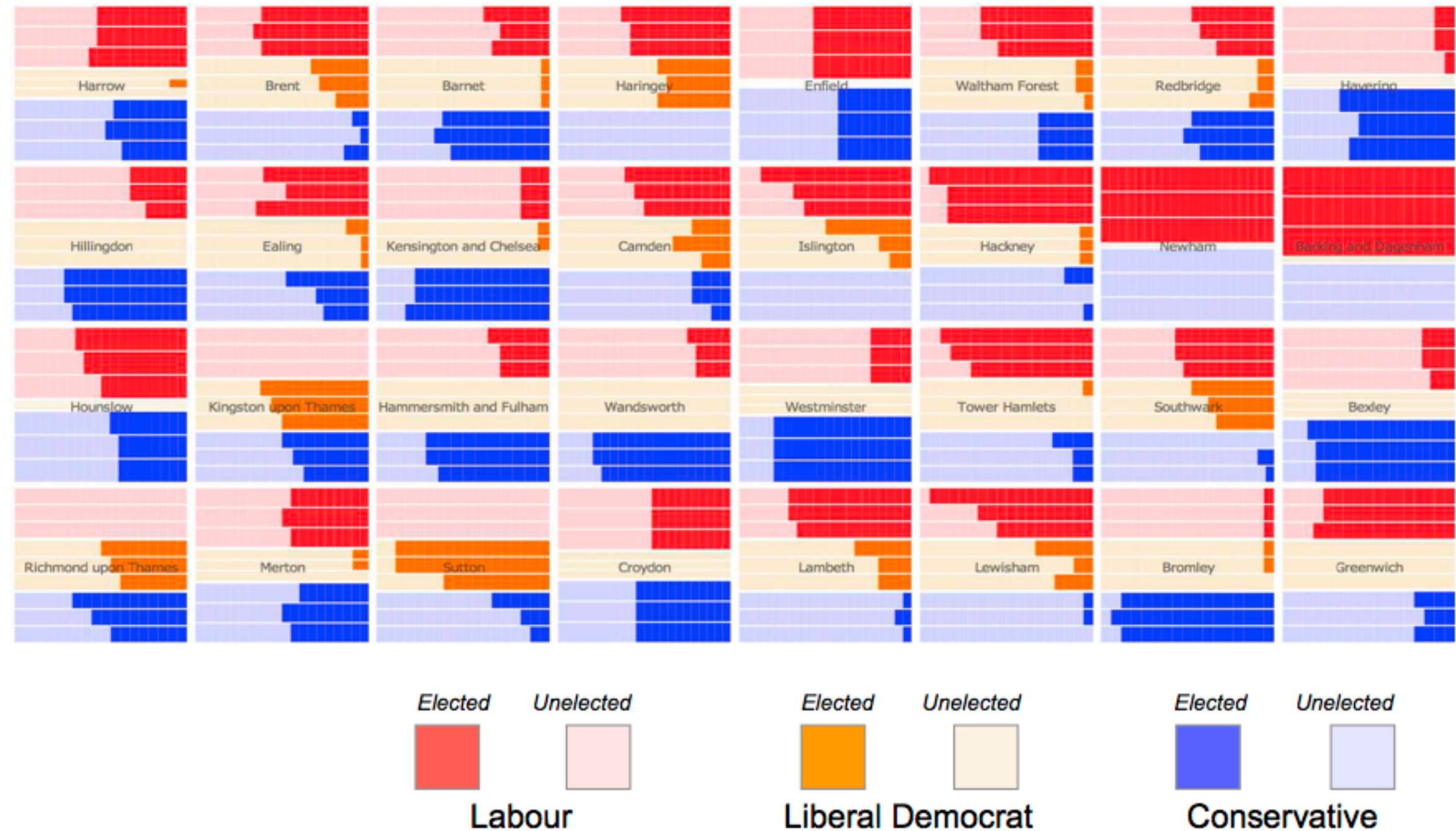
Derived Data

BallotMaps

- ballots in the UK are alphabetically ordered
 - govt: not sufficient to affect electoral outcome
 - vis researcher hunch: it matters!
- vis project
 - task: compare geographic regions of voting and spatial position of candidate name on ballot paper
 - data: Greater London elections 2010
 - geographic location, candidate name, alphabetical position in ballot, # candidate votes, party, elected/lost
 - color coding alone will not save the day!
 - derived new data
 - alphabetical position within the party
 - vote order within party

BallotMaps: deriving data

- bias exists in regions where systematic structure in bar lengths visible
 - yes in some
 - no in others



[BallotMaps: Detecting name bias in alphabetically ordered ballot papers. Wood, Badawood, Dykes, Slingsby. *IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 2011)*, 17(12): 2384-2381, 2011]

Four strategies to handle complexity

→ *Derive*



- derive new data to show within view
- change view over time
- facet across multiple views
- reduce items/attributes within single view

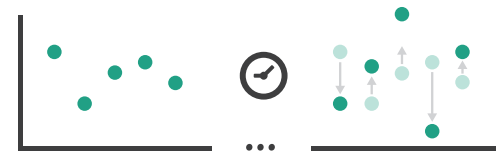
more at:

Visualization Analysis and Design.

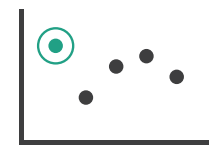
Munzner.AK Peters Visualization Series, CRC Press, 2014.

Manipulate

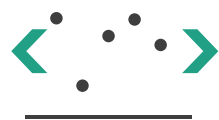
→ Change



→ Select

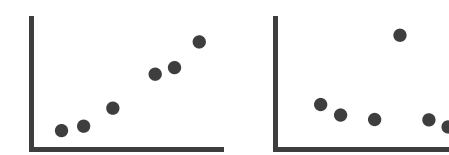


→ Navigate

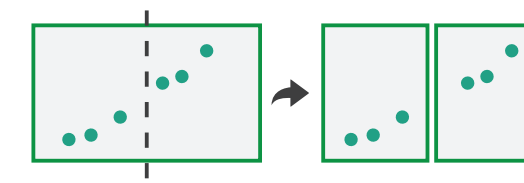


Facet

→ Juxtapose



→ Partition

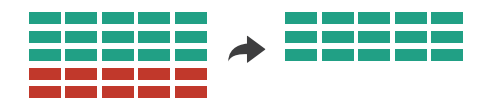


→ Superimpose

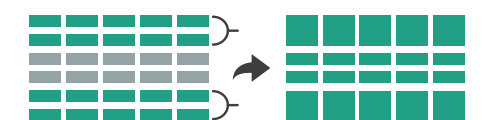


Reduce

→ Filter



→ Aggregate



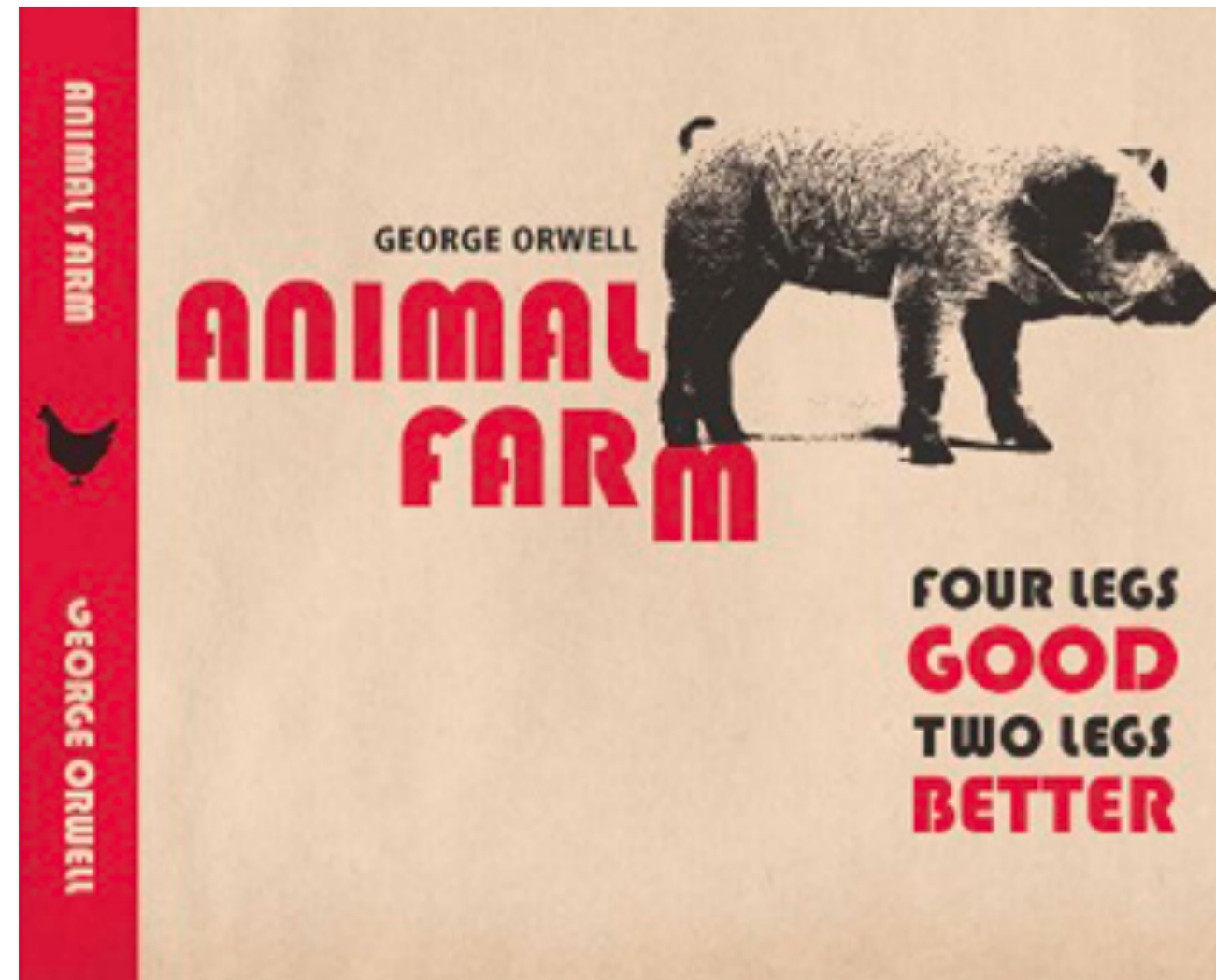
→ Embed



**Vignette 4:
Difficulties of Depth
(Another Cautionary Tale)**

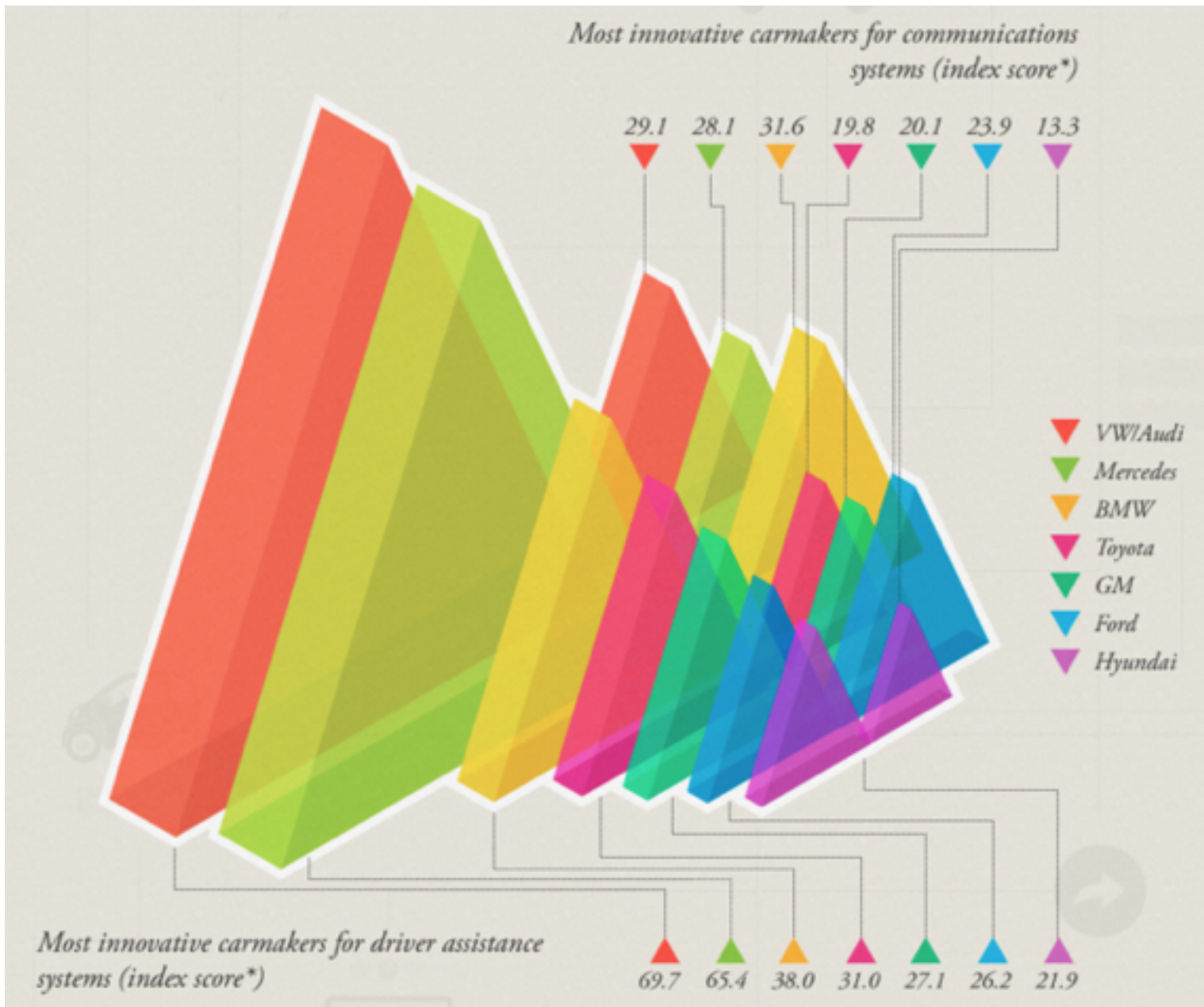
Visual encoding: 2D vs 3D

- 2D good, 3D better?
 - not so fast...

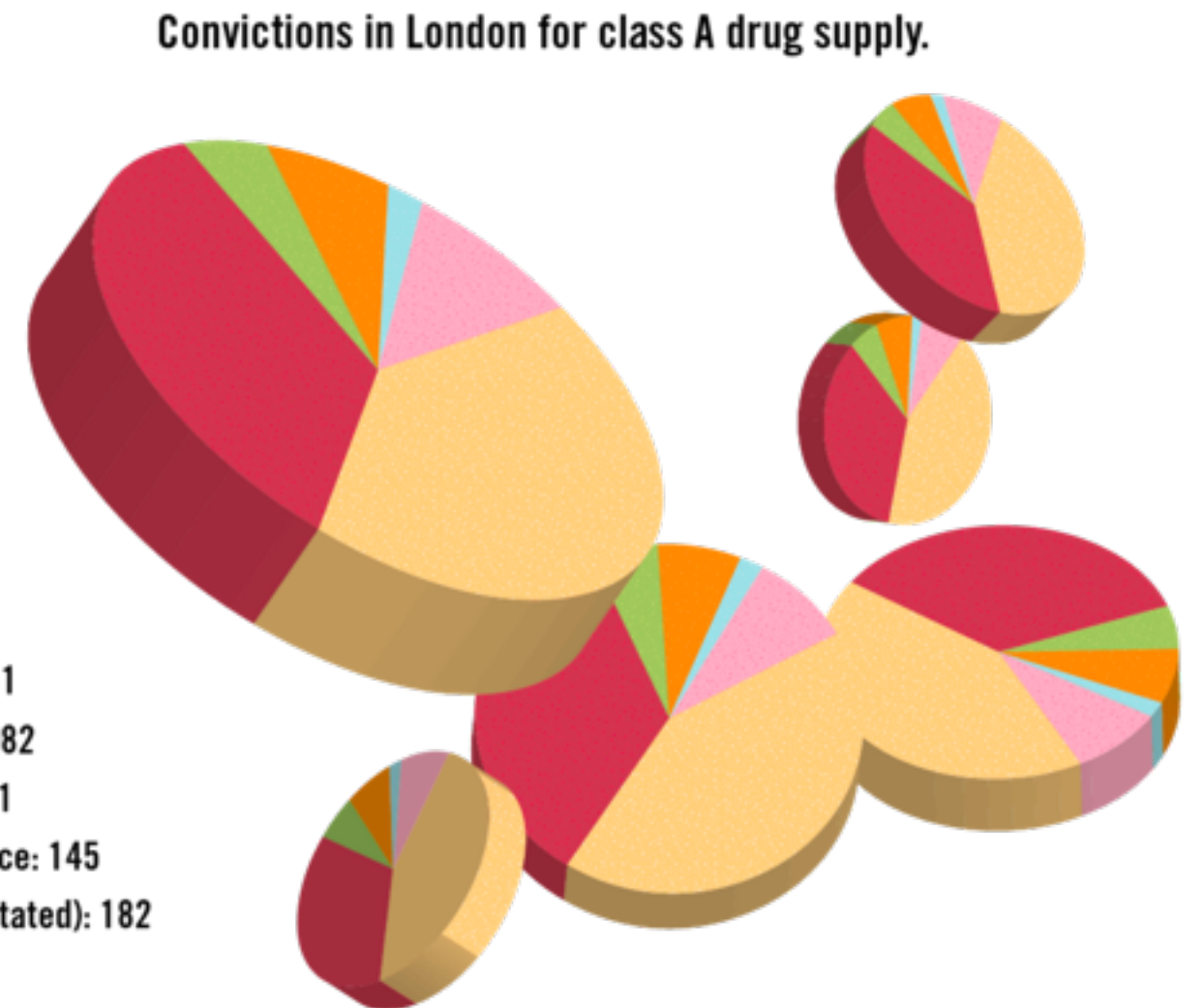


<http://amberleyromo.com/images/Bookcover/Animal-Farm.png>

Unjustified 3D all too common, in the news and elsewhere



<http://viz.wtf/post/137826497077/eye-popping-3d-triangles>

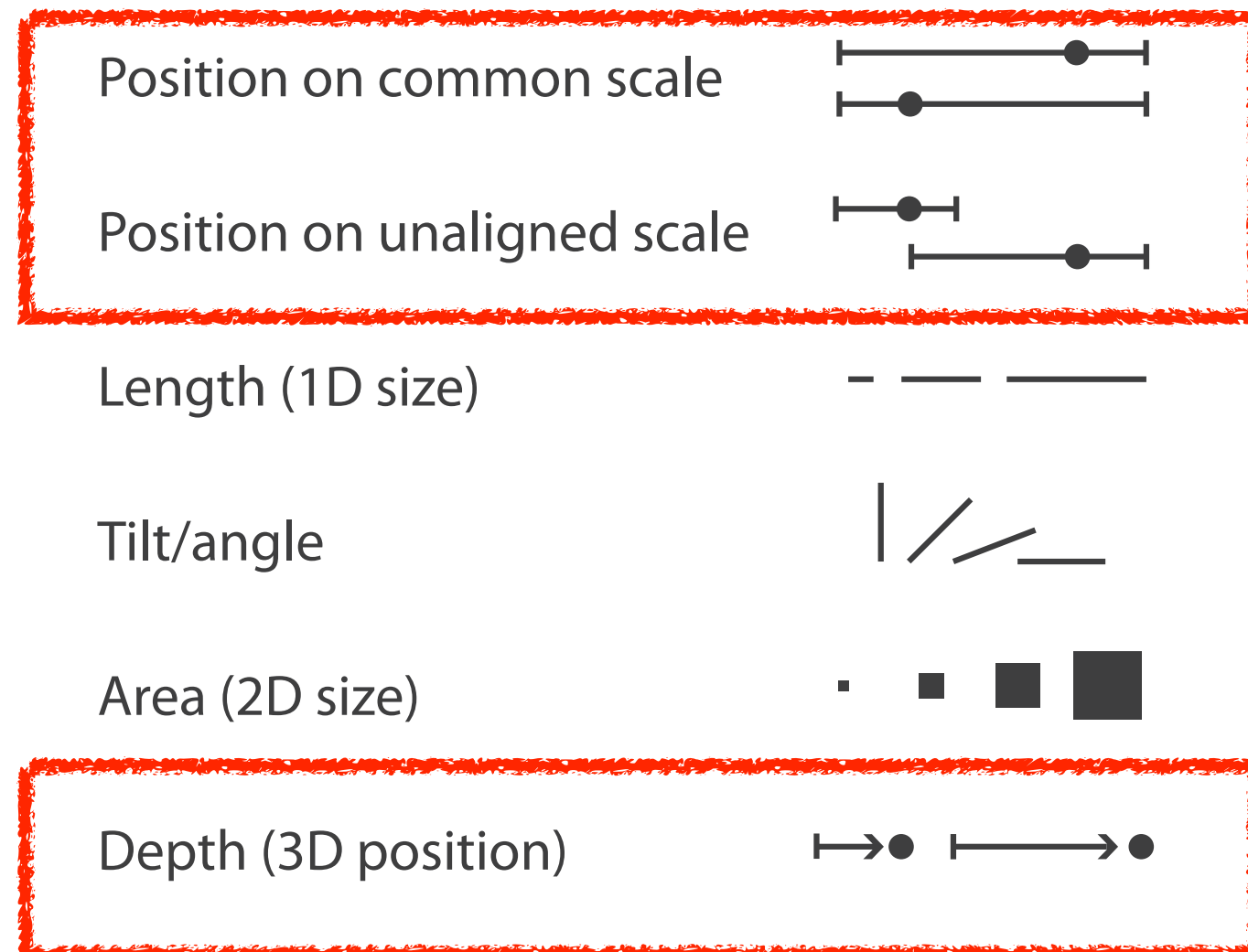


<http://viz.wtf/post/139002022202/designer-drugs-ht-ducqn>

Depth vs power of the plane

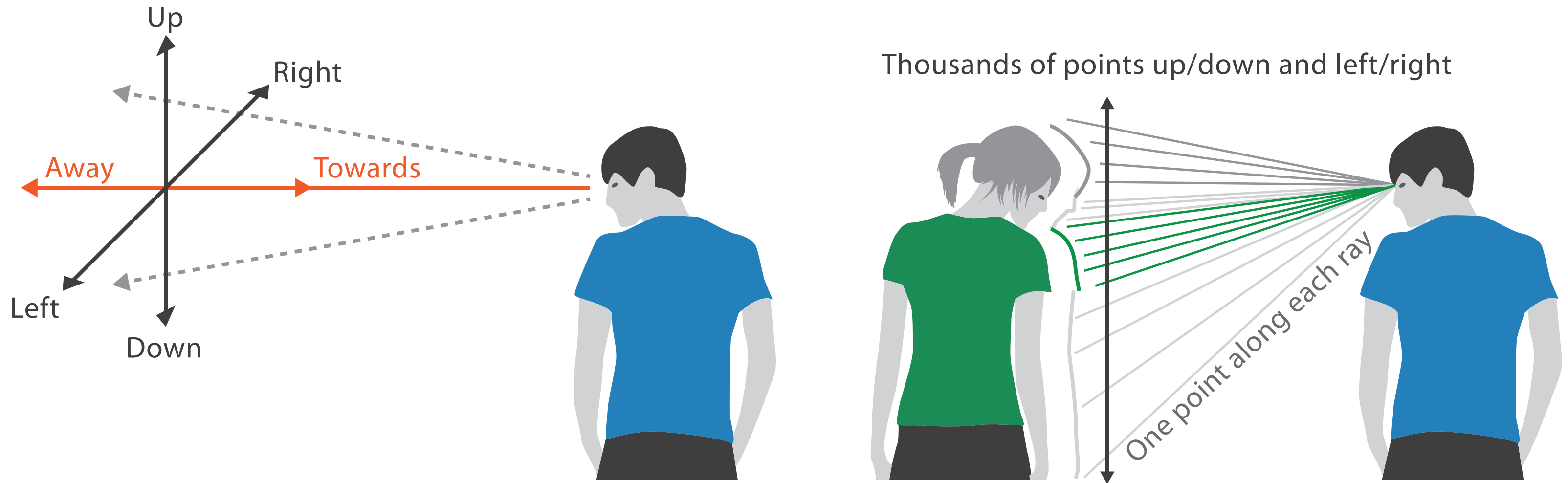
- high-ranked spatial position channels: **planar** spatial position
 - not depth!

→ Magnitude Channels: **Ordered** Attributes



Life in 3D?...

- we don't really live in 3D: we **see** in 2.05D
 - acquire more info on image plane quickly from eye movements
 - acquire more info for depth slower, from head/body motion

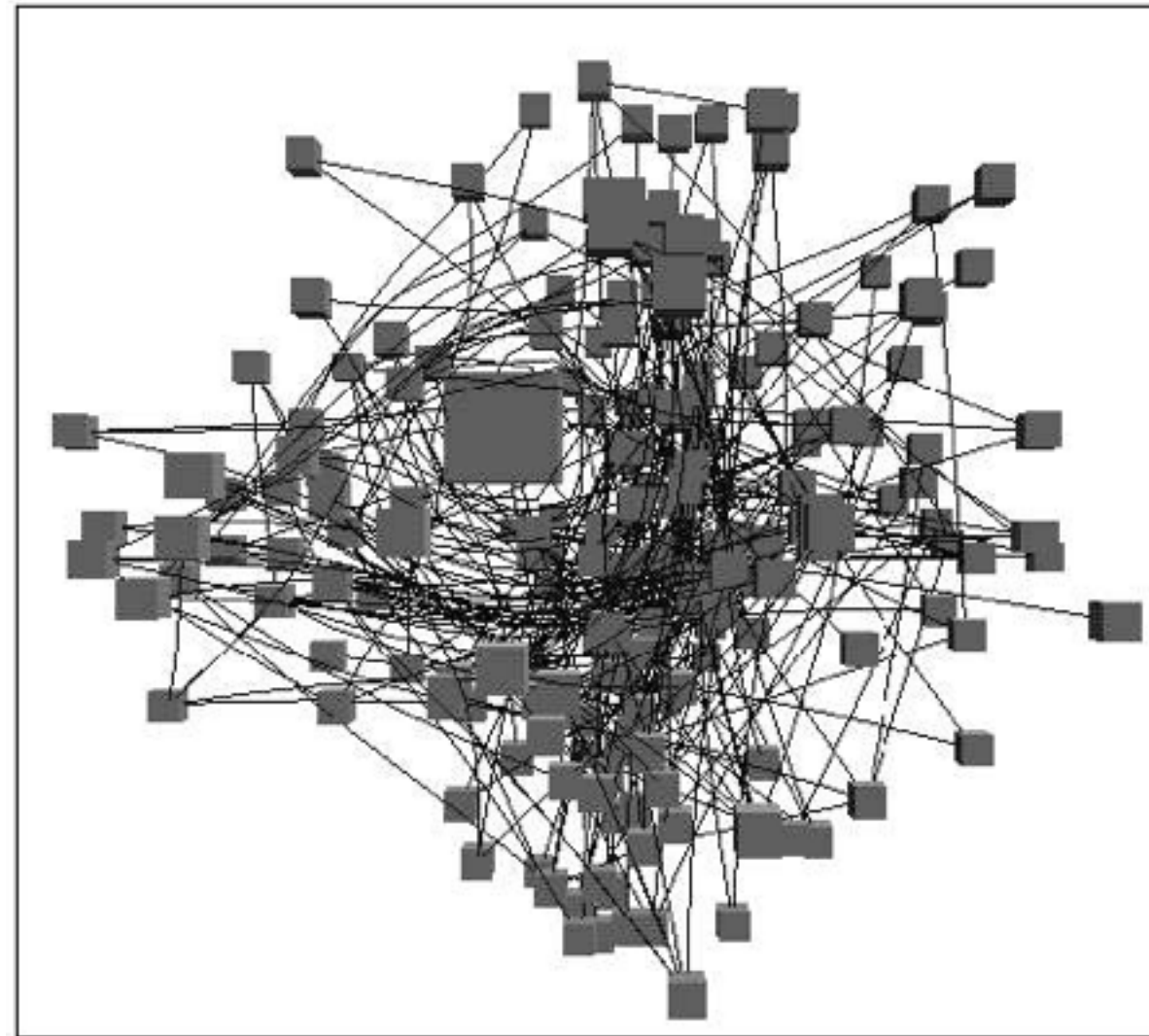


[adapted from *Visual Thinking for Design*. Ware. Morgan Kaufmann 2010.]

We can only see the outside shell of the world

Occlusion hides information

- occlusion
- interaction complexity



[Distortion Viewing Techniques for 3D Data. Carpendale et al. InfoVis 1996.]

Perspective distortion loses information

- perspective distortion
 - interferes with all size channel encodings
 - power of the plane is lost!



*[Visualizing the Results of Multimedia Web Search Engines.
Mukherjea, Hirata, and Hara. InfoVis 96]*

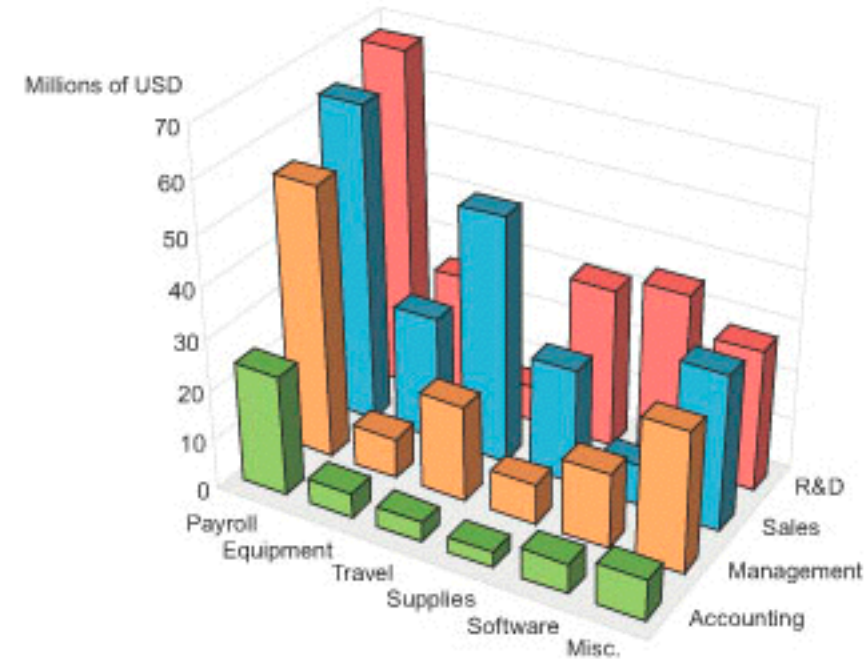
3D vs 2D bar charts

- 3D bars never a good idea!

Graph Design I.Q. Test

Question 7: Which graph makes it easier to determine R&D's travel expense?

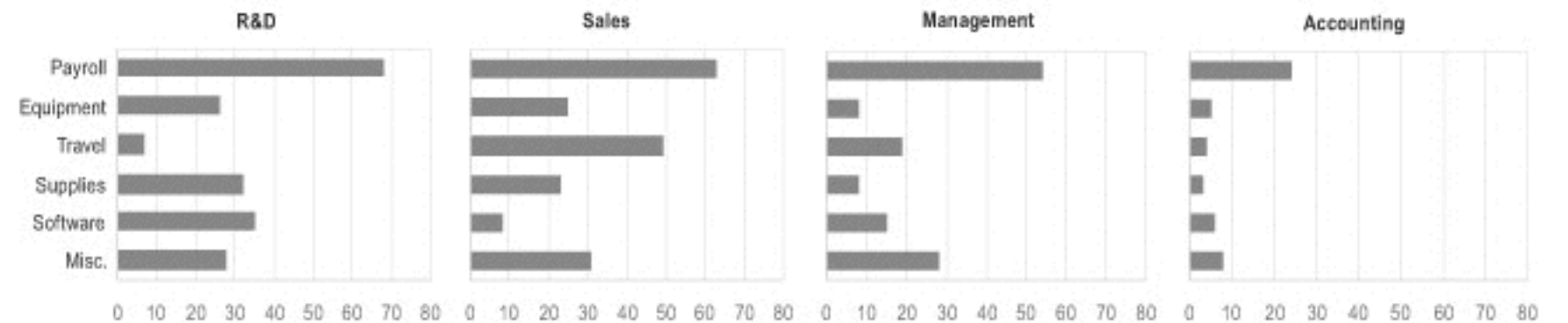
2006 Expenses by Department



3-D Bar Graph (left)

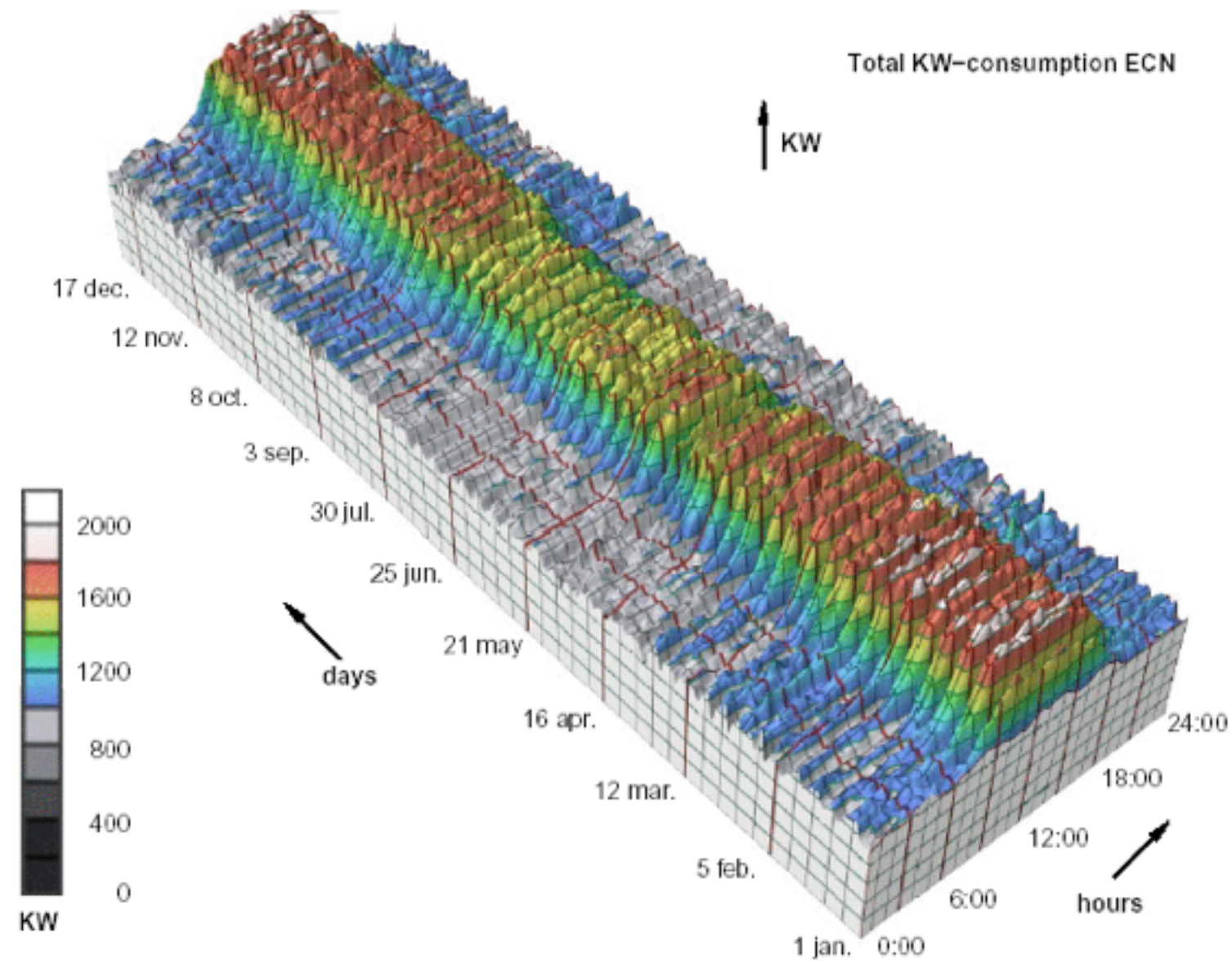
2-D Bar Graphs (below)

2006 Expenses by Department in Millions of USD



No unjustified 3D example: Time-series data

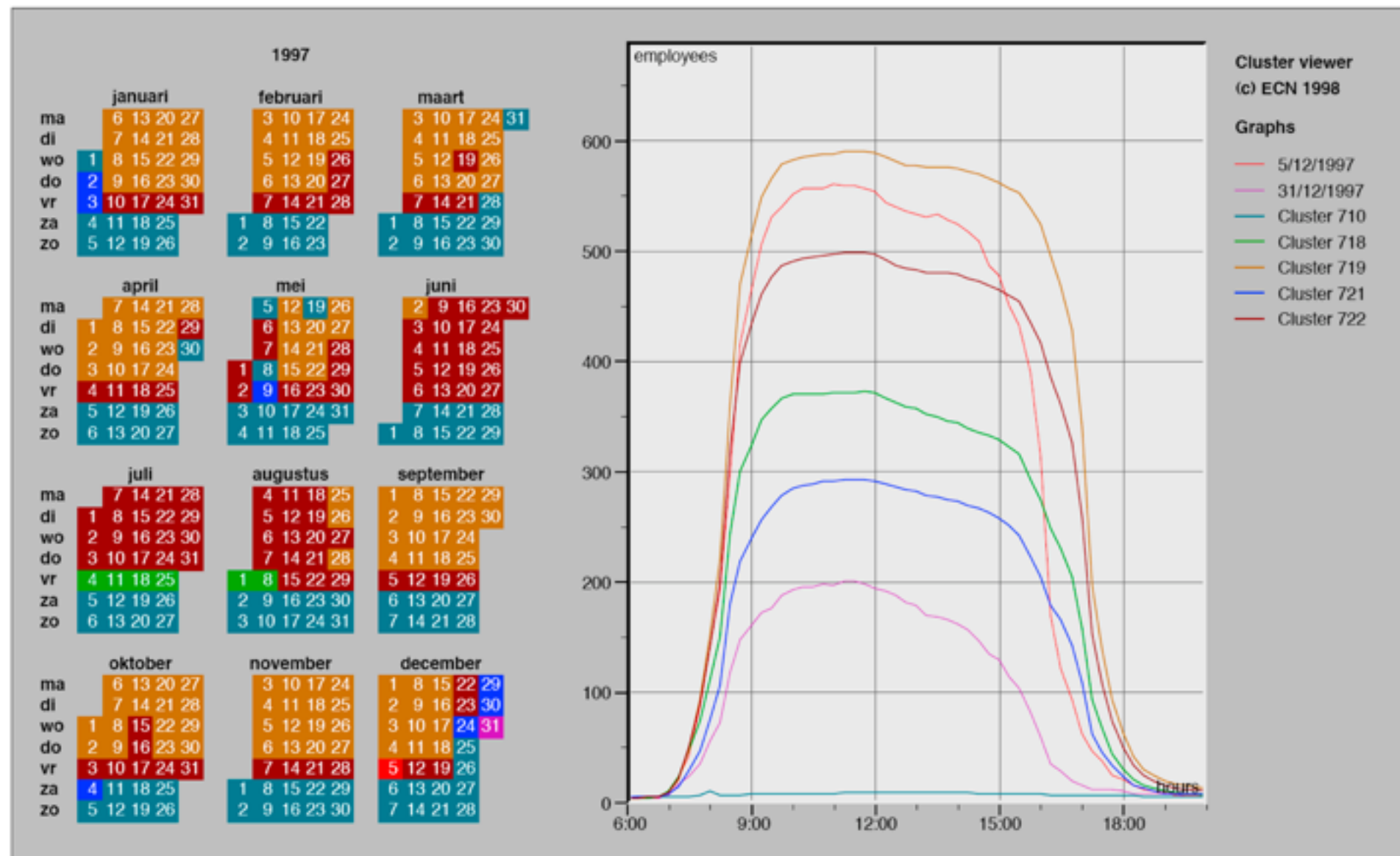
- extruded curves: detailed comparisons impossible



[Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc. InfoVis 99.]

No unjustified 3D example: Transform for new data abstraction

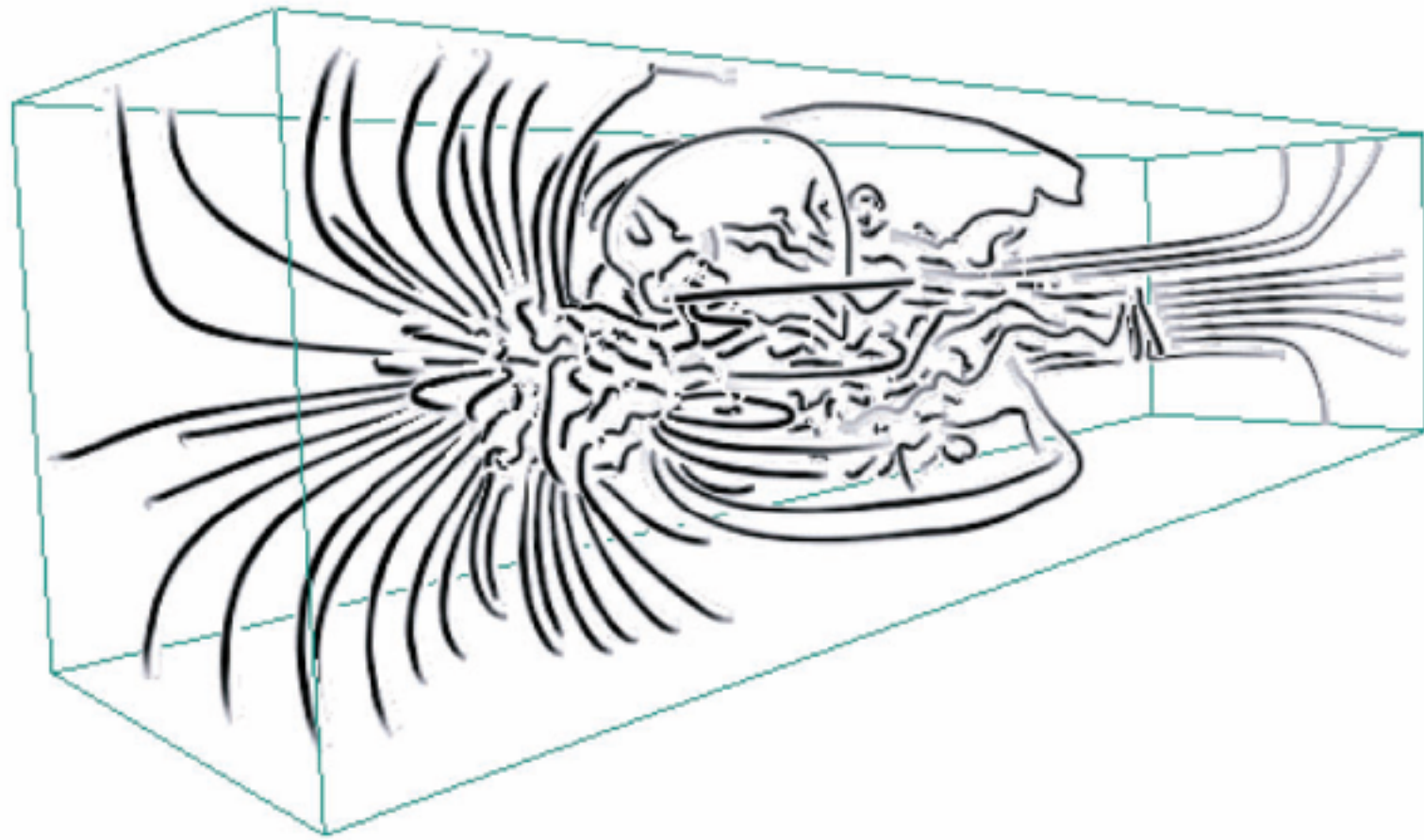
- derived data: cluster hierarchy
- juxtapose multiple views: calendar, superimposed 2D curves



[Cluster and Calendar based Visualization of Time Series Data. van Wijk and van Selow, Proc. InfoVis 99.]

Justified 3D: shape perception

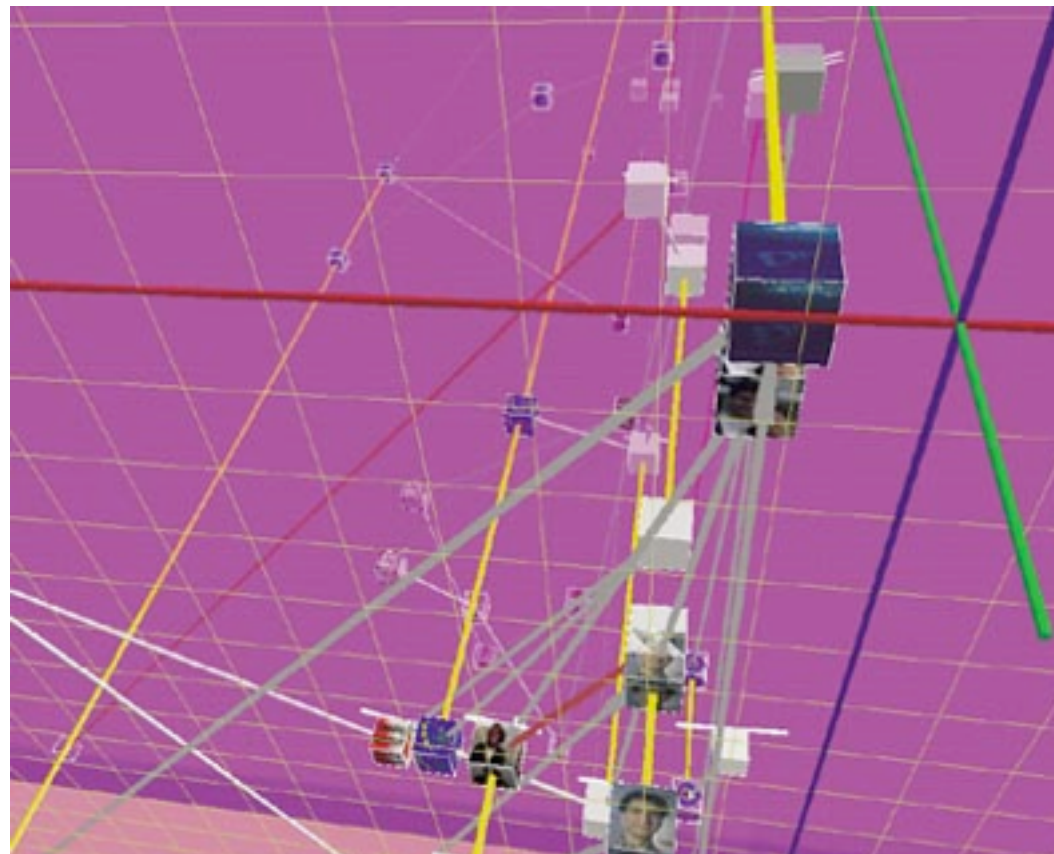
- benefits outweigh costs when task is shape perception for 3D spatial data
 - interactive navigation supports synthesis across many viewpoints



[Image-Based Streamline Generation and Rendering. Li and Shen. IEEE Trans. Visualization and Computer Graphics (TVCG) 13:3 (2007), 630–640.]

No unjustified 3D

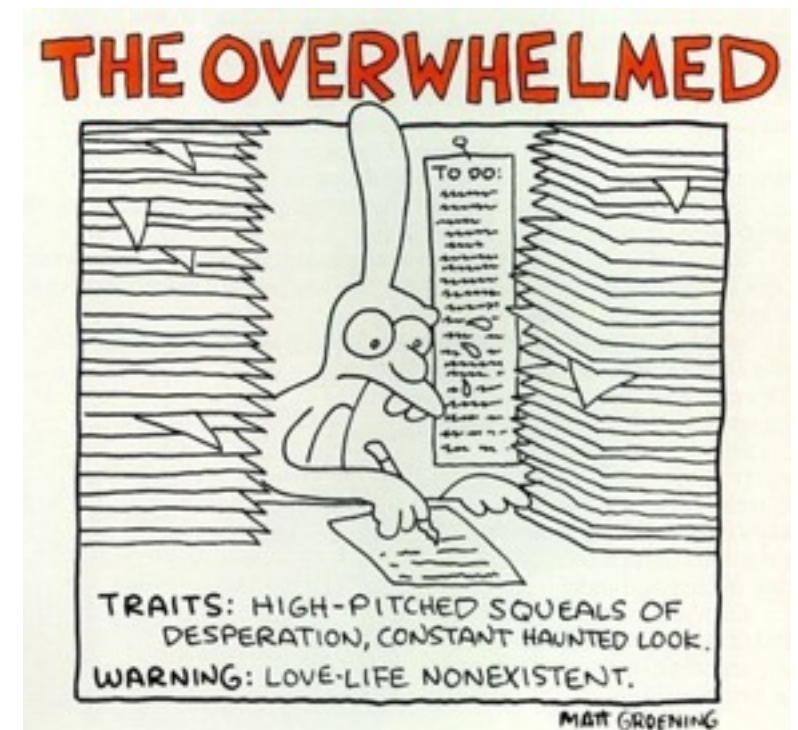
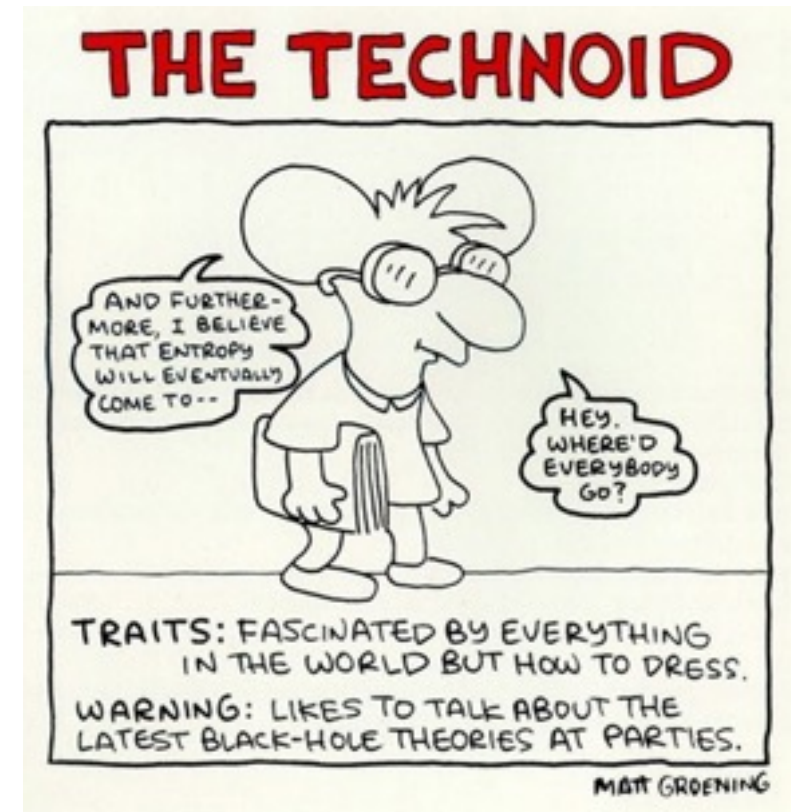
- 3D legitimate for true 3D spatial data
- 3D needs very careful justification **for abstract data**
 - enthusiasm in 1990s, but now skepticism
 - be especially careful with 3D for point clouds or networks



[WEBPATH-a three dimensional Web history. Frecon and Smith. Proc. InfoVis 1999]

Wrap-up

- a tale of two tools
 - exploration: Overview
 - collaboration between CS and journalism: methods & rewards
 - reasoning about four levels of vis design
 - presentation: TimeLineCurator
 - visual curation of imperfect computational results
 - the importance of being brisk: speedup vs eureka moment
- two cautionary tales
 - guidance on color & 3D from vis literature



More Information

[@tamaramunzner](https://twitter.com/tamaramunzner)

- this talk

www.cs.ubc.ca/~tmm/talks.html#cj16

- book

<http://www.cs.ubc.ca/~tmm/vadbook>

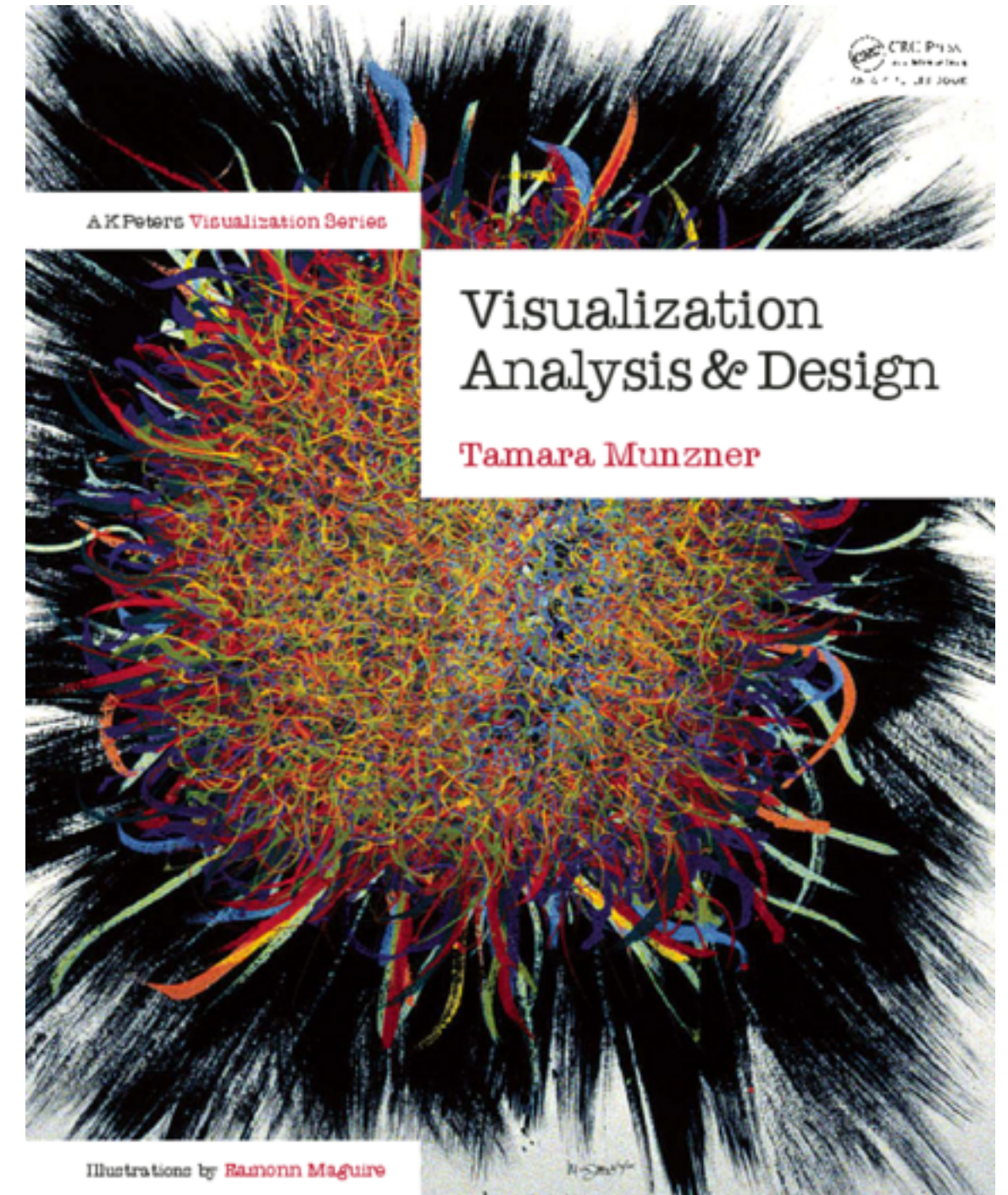
– 20% off promo code, book+ebook combo: HVN17

– <http://www.crcpress.com/product/isbn/9781466508910>

- papers, videos, software, talks, courses

<http://www.cs.ubc.ca/group/infovis>

<http://www.cs.ubc.ca/~tmm>



Visualization Analysis and Design.
Munzner. A K Peters Visualization Series, CRC Press, Visualization Series, 2014.