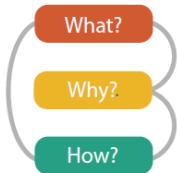


Information Visualization and Visual Analytics (M1522.000500)

Definitions and Motivations

Jinwook Seo, Ph. D.

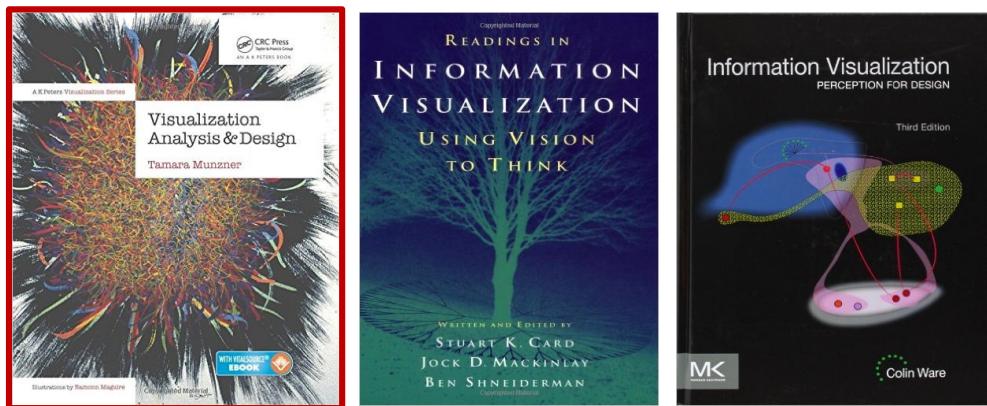
Professor, Dept. of Computer Science and Engineering
Seoul National University



The Big Picture

Definition of Visualization (VIS)

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

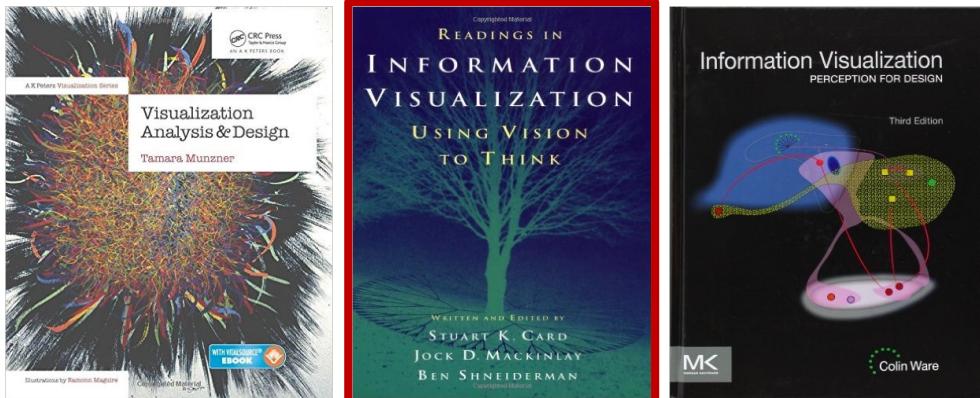


Definition of Visualization (VIS)

The use of computer-supported, interactive, visual representations of

abstract data to amplify **cognition**

- Stuart Card, Jock Mackinlay, Ben Shneiderman, 1999

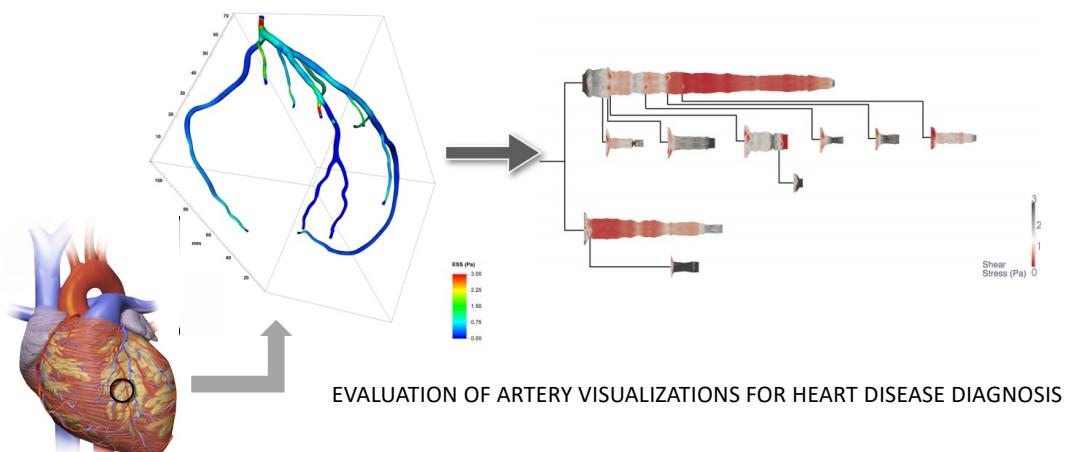


Information Visualization and Visual Analytics – Definitions and Motivations

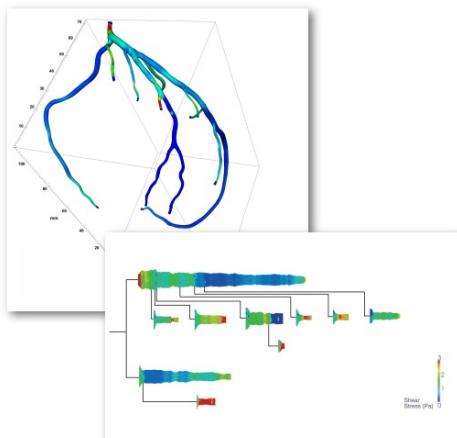
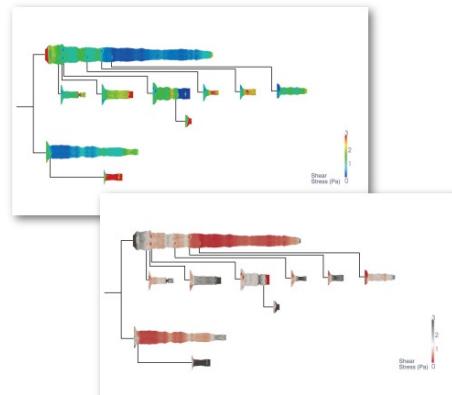
InfoVis vs. SciVis

InfoVis vs. SciVis

- Spatial representations are **given** or have to be **designed**



Even for Physical Objects

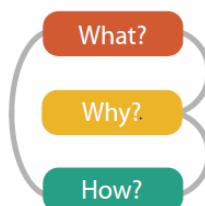
3D vs. 2D**rainbow vs. diverging**

<https://www.cs.ubc.ca/~tmm/courses/547-17/slides/borkin-artery-subset-slides.pdf>

Information Visualization and Visual Analytics – Definitions and Motivations

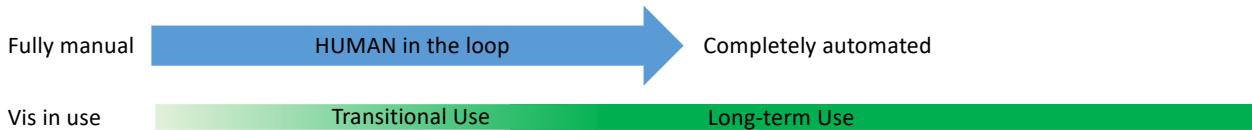
The Big Picture**The Big Picture**

- VIS is suitable when there is a need to **augment human capabilities**
 - vs. replace people with computational decision-making methods
 - HUMAN in the decision-making loop!
- **DESIGN** visual representations to help people perform **tasks** more **effectively**
- Design space is **HUGE!**
 - How to create visual representations
 - How to interact with visual representations
 - Effectiveness of a design
- Resource limitations
 - Computers, Humans, Displays
- Analysis Instance
 - WHY the **user** needs it, WHAT **data** is shown, HOW the **idiom** is designed



Information Visualization and Visual Analytics – Definitions and Motivations

Human in the loop



- Vis allows people to analyze data **when they don't know exactly what questions they need to ask in advance**
 - vs. purely computational techniques for well-defined questions (e.g., high-frequency trading)
- Visualization systems are appropriate for use
 - when your goal is to **augment human capabilities**
 - when the problem is **ill-specified** (i.e., people don't know how to approach it)

Transitional Use

- “work itself out of a job”
 - helping designers of purely computational solutions
 - once computational solutions are in place, vis system retires
- VIS tool in the **early** stages of a transition (highly exploratory)
 - gain a clearer understanding of user’s task and analysis requirements
 - before starting to develop formal mathematical/computational models
- VIS tool in the **middle** stages of a transition
 - for designers of a purely computational solution
 - e.g., tuning parameters of high-frequency trading, tuning hyperparameters of DL models

Why Have a Human in the Loop?

Transitional Use

- VIS tool in the late stages of a transition
 - help end users (or designers) make deployment decisions of whether the automatic system is doing the right thing



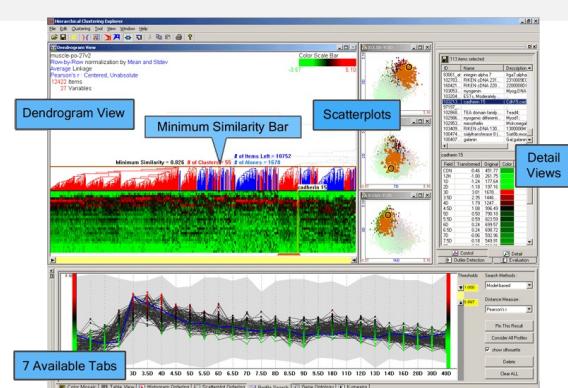
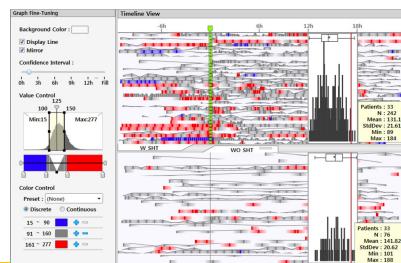
- vis tool abandoned after that decision is made
 - vs.
- vis tool in play with long-term use to **monitor** a system

Information Visualization and Visual Analytics – Definitions and Motivations

Why Have a Human in the Loop?

Long-lasting Use – Exploratory Visualization Tools

- Many analysis problems are ill specified
 - people don't know how to approach the problem.
- Design vis tools for long-term use
 - A person will stay in the loop indefinitely
- Exploratory analysis for (scientific) discovery
 - help users generate and check [hypotheses](#)



Hierarchical Clustering Explorer

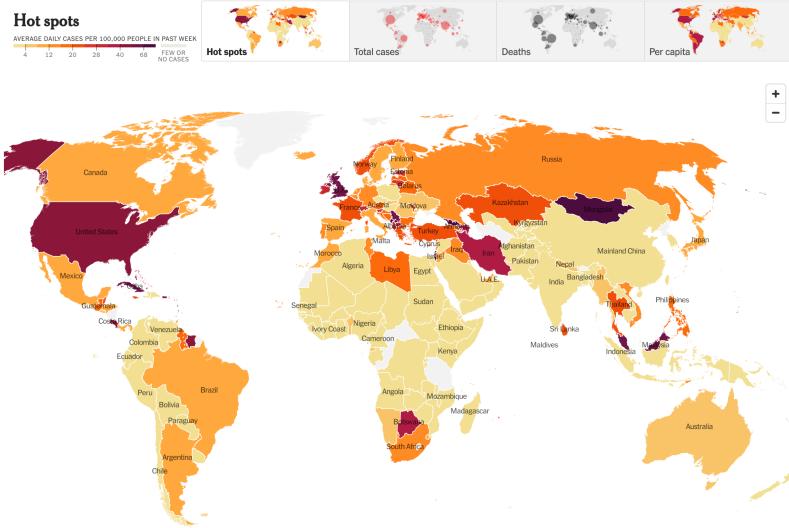
StroscopE

Information Visualization and Visual Analytics – Definitions and Motivations

Why Have a Human in the Loop?

Long-lasting Use – Presentation

- Design vis tools for presentation
 - storytelling
 - explain something you already know



<https://www.nytimes.com/interactive/2021/world/covid-cases.html>

Information Visualization and Visual Analytics – Definitions and Motivations

Why Have a Human in the Loop?

Long-lasting Use – Presentation

Published: February 2, 2010

Budget Forecasts, Compared With Reality

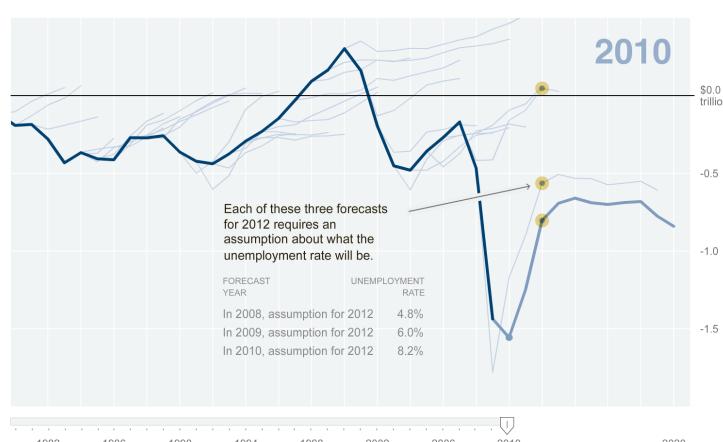
Just two years ago, surpluses were predicted by 2012. How accurate have past White House budget forecasts been?

1 2 3 4 5 6 NEXT >

Reasons for error

Budget forecasts require assumptions about both policy and the economy, and both can turn out to be spectacularly wrong.

In 2008, for example, the unemployment rate in 2012 was assumed to be 4.8 percent. The most recent assumption is 8.2 percent.



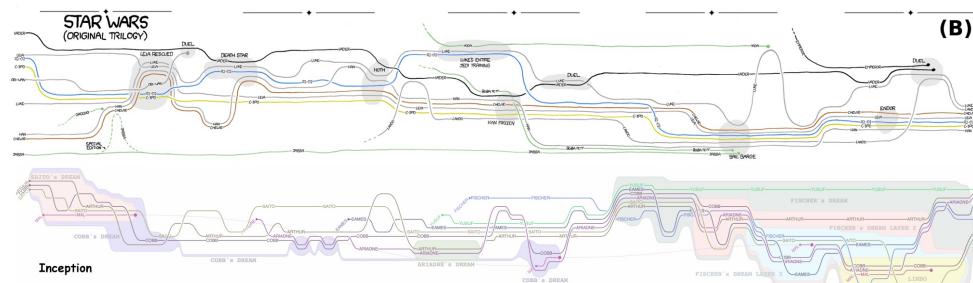
<https://archive.nytimes.com/www.nytimes.com/interactive/2010/02/02/us/politics/20100201-budget-porcupine-graphic.html?r=0>

Information Visualization and Visual Analytics – Definitions and Motivations

Why Have a Computer in the Loop?

Why Have a Computer in the Loop?

- The scope of what people are willing and able to do manually is strongly limited by their **attention span**
- Scalability & Generalizability
- Datasets changing dynamically over time



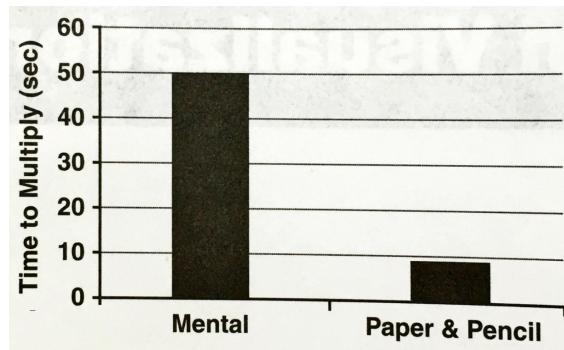
Information Visualization and Visual Analytics – Definitions and Motivations

Why Use an External Representation?

External Representations (External Memory)

- Finding the *artificial memory* that best supports our natural means of perception
 - Bertin, 1983

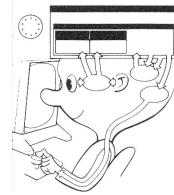
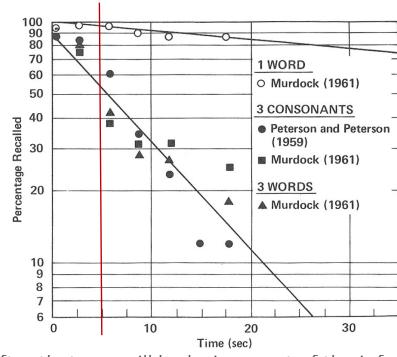
$$\begin{array}{r} 34 \\ \times 72 \\ \hline 68 \\ 238 \\ \hline 2448 \end{array}$$



Information Visualization and Visual Analytics – Definitions and Motivations

Human Information Processor

- The working memory is your register set
- Access in chunks
 - Task dependent construct
 - 7 ± 2 (Miller)
- working memory decay with time
 - Content dependent
 - Limited attention span
 - 5s attention span seems like a good idea, since after that you will be losing most of the information
 - commit to longer term memory
 - “external cognition tool” like writing on a piece of paper



Information Visualization and Visual Analytics – Definitions and Motivations

Why Use an External Representation?

External Representations (External Memory)

- Surpass limitations of our own internal cognition and memory
- Offload cognition and memory usage to the perceptual system
- Visualization (e.g., Diagrams) for perceptual inference
 - Organized information by spatial location
 - Recall vs. Recognition
 - Accelerating both search and recognition by grouping items/information

Information Visualization and Visual Analytics – Definitions and Motivations

Why Depend on Vision?

- Visualization is based on exploiting the human visual system
- Visual system: very high-bandwidth channel to brain
 - Vision: 100 MB/s
 - Ears: <100 b/s
- Preattentive processing
 - visual information processing occurs in parallel at the preconscious level
 - visual **popout**
- Sonification, Touchification, Smellification, Tastification...

Information Visualization and Visual Analytics – Definitions and Motivations

How many sevens?

2398419309213985874506209348952034809502
3984210293841909238740129384610329849238
4265293845013945594858601239480234958728
4596394058640598239485802394895029348658
4561024596234851604569828309458673049561
3045916459086130495298646658956405196809
5866304598683049561835601830459680345907
6283486510465183560241620945613045618304
5968230459630459860395620349568204385362

Slide Idea from Colin Ware

Information Visualization and Visual Analytics – Definitions and Motivations

How many sevens?

Color Makes Them Pop Out

2398419309213985874506209348952034809502
3984210293841909238740129384610329849238
4265293845013945594858601239480234958728
4596394058640598239485802394895029348658
4561024596234851604569828309458673049561
3045916459086130495298646658956405196809
5866304598683049561835601830459680345907
6283486510465183560241620945613045618304
5968230459630459860395620349568204385362

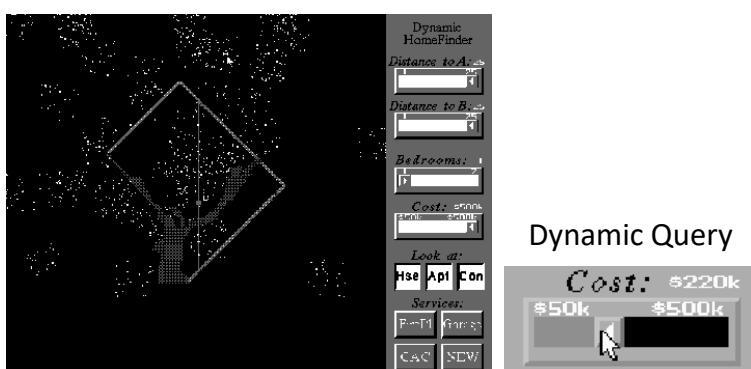
Slide Idea from Colin Ware

Information Visualization and Visual Analytics – Definitions and Motivations

Why use Interactivity?

Why use Interactivity?

- Impossible to show everything at once for a LARGE dataset
- Interactivity for handling complexity and volume
 - increasing the scope and capabilities of vis tools



Dynamic Query

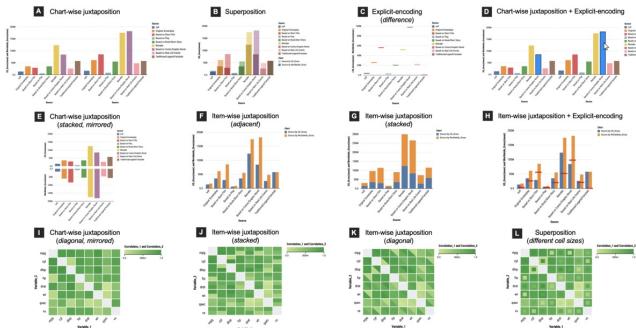
<https://youtu.be/5X8XY9430fM>

Information Visualization and Visual Analytics – Definitions and Motivations

Why is the Vis Idiom Design Space Huge?

Visualization Idioms

- distinct approach to creating and manipulating visual representations
- many ways to create a visual encoding of data



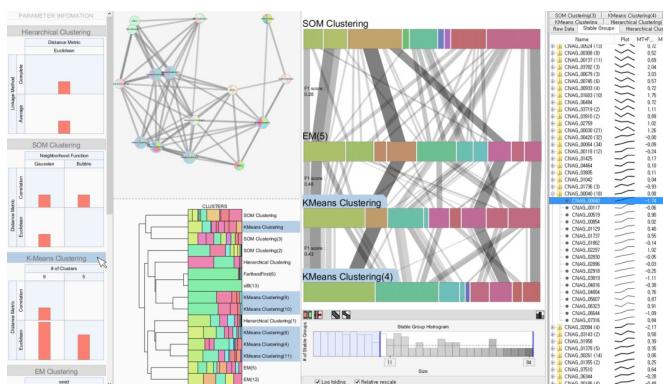
S. LYI, J. JO AND J. SEO, "Comparative Layouts Revisited: Design Space, Guidelines, and Future Directions" in *IEEE Transactions on Visualization & Computer Graphics*, vol. 27, no. 02, pp. 1525-1535, 2021. doi: 10.1109/TVCG.2020.3030419

- framework for thinking about the space of vis design idioms systematically

Why is the Vis Idiom Design Space Huge?

Why Is the Vis Idiom Design Space Huge?

- complicated idioms: link together multiple simple charts through interaction
- many more design choices

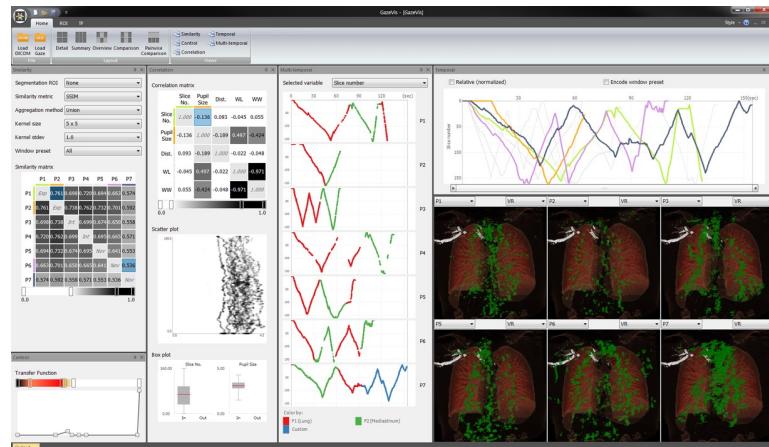


L'YI S, KO B, SHIN D, CHO YJ, LEE J, KIM B, SEO J. XCluSim: a visual analytics tool for interactively comparing multiple clustering results of bioinformatics data. BMC Bioinformatics. 2015;16 Suppl 11(Suppl 11):S5. doi: 10.1186/1471-2105-16-S11-S5. Epub 2015 Aug 13. PMID: 26328893; PMCID: PMC4547151.

Why is the Vis Idiom Design Space Huge?

Why Is the Vis Idiom Design Space Huge?

- complicated idioms: link together multiple simple charts through interaction
 - many more design choices



Song H, Lee J, Kim TJ, Lee KH, Kim B, Seo J. GazeDx: Interactive Visual Analytics Framework for Comparative Gaze Analysis with Volumetric Medical Images. *IEEE Trans Vis Comput Graph.* 2017 Jan;23(1):311-320. doi: 10.1109/TVCG.2016.2598796. PMID: 27875147.

Information Visualization and Visual Analytics – Definitions and Motivations

Why Focus on Tasks?

Why Focus on Tasks?

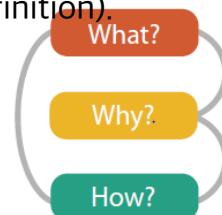
- A tool that serves well for one task can be poorly suited for another, for exactly the same dataset.
 - Reframing the users' task from domain-specific form into abstract form
 - "contrast the prognosis of patients who were intubated in the ICU more than one month after exposure to patients hospitalized within the first week"
 - "see if the results for the tissue samples treated with LL-37 match up with the ones without the peptide"

Information Visualization and Visual Analytics – Definitions and Motivations

Why Focus on Tasks?

- A tool that serves well for one task can be poorly suited for another, for exactly the same dataset.
- Reframing the users' task from domain-specific form into abstract form
 - "contrast the prognosis of patients who were intubated in the ICU more than one month after exposure to patients hospitalized within the first week"
 - "see if the results for the tissue samples treated with LL-37 match up with the ones without the peptide"
 - ➔ "compare values between two groups"

두개의 그룹을 대표하는 값들을 비교하는 것 (abstract task definition).
Domain independent하게 정의가능
- Analysis framework provide generic words to describe why



Effectiveness

- A visualization is more **effective** than another visualization if the information conveyed by one visualization is **more readily perceived** than the information in the other visualization.
 - Goal of vis is to support user tasks → effectiveness
 - correctness, accuracy, and truth
 - "it's not just about making pretty pictures"
 - beautiful but not effective?
 - Any **visualization is an abstraction** where choices are made about **which aspects to emphasize**.
- 유저가 더 쉽게 정보를 접할 수 있을 때 좋은
visualization. 노력을 덜 넣을 때
- Distortion이 있을 수 밖에 없는데 유저에게 방해가 안되는.

Why Are Most Designs Ineffective?

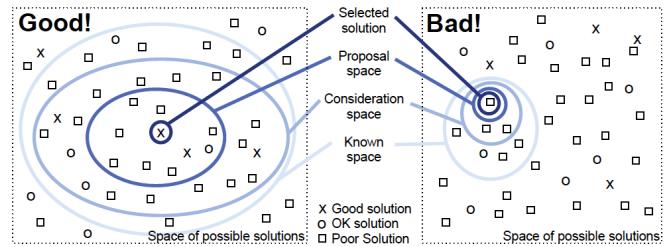
Why Are Most Designs Ineffective? Huge Design Space!

Not to optimize (to find the very best choice),

but to **satisfy** (to find one of the many possible good solutions rather than one of the even larger number of bad ones)

- good vs. bad strategy to traverse the design space
 - the known and consideration spaces are large vs.
 - the spaces are small

the holy grail of design:

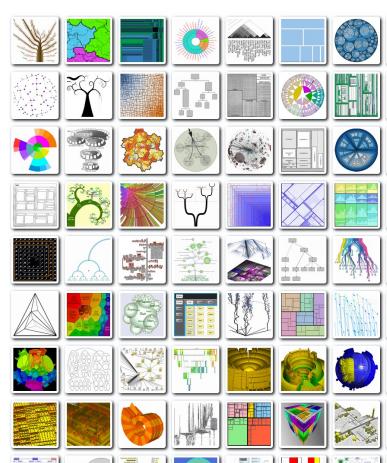


- consider multiple alternatives in parallel and then choose the best

Why is Validation Difficult?

Why is Validation Difficult?

- SO MANY QUESTIONS that you could ask when considering whether a vis tool has met your design goals
- better, faster, effective, insight, engagement, familiar, accurate, satisfactory...
- when to use what evaluation methods for validating vis designs?



Why Are There Resource Limitations?

- computational capacity
- human perceptual and cognitive capacity
- display capacity
- Scalability is a central concern
- Computational capacity (computer time and memory)
 - Soft and hard constraints on the availability of these resources.
 - Computational complexity of algorithms for dataset preprocessing, transformation, layout, and rendering is a major concern.

Information Visualization and Visual Analytics – Definitions and Motivations

Human Perceptual and Cognitive Capacity

- Human memory for things that are not directly visible is notoriously limited
- Change blindness (selective attention) **유저가 한가지만 보고있는것**
 - even very large changes are not noticed if we are attending to something else in our view
 - store surprisingly little information internally in visual working memory



copyright (c) 1999 Daniel J. Simons. All rights reserved.

<https://youtu.be/vJG698U2Mvo>

Information Visualization and Visual Analytics – Definitions and Motivations

Change Blindness

- Change Blindness

- when we fail to notice drastic changes if our attention is directed elsewhere

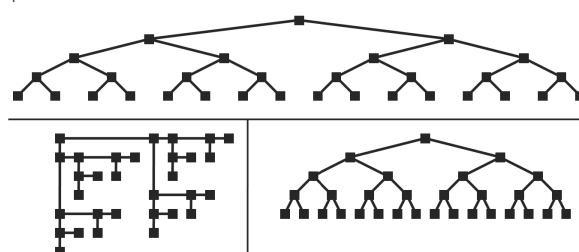


→ hard to track complex and widespread changes in multi-frame animations

Why are there Resource Limitations?

Display Capacity

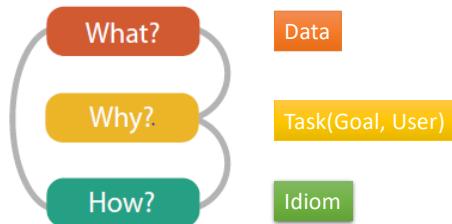
- The resolution of the screen is not enough to show all desired information simultaneously
- Information density :
 - a measure of the amount of information encoded
 - vs. the amount of unused space.



- Minimizing the need for navigation and exploration
 - vs. Costs of showing too much at once

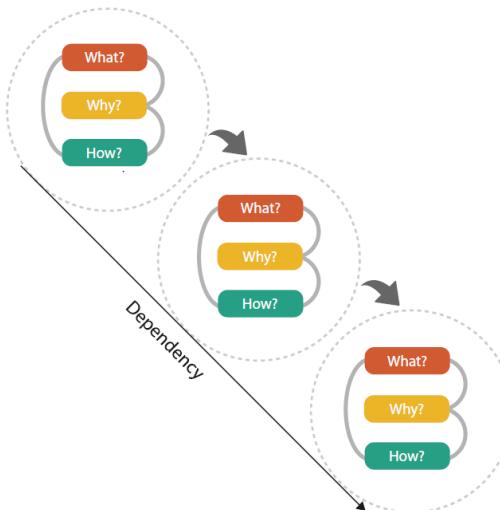
Why Analyze?

- Analyze existing systems → designing new ones
- **Three-part Analysis Framework**
 - **What** data the user sees
 - **Why** the user intends to use a vis tool
 - **How** the visual encoding and interaction idioms are constructed in terms of design choices



How to Analyze Complex Vis Tools

- One analysis instance for each task
- a sequence of analysis instances chained together
 - express dependencies
 - output of a prior instance serves as the input to a subsequent one



- What's Vis and Why Do It?
- Huge Design Space
- Resource Limitation
- Analysis Framework
- Questions?