Paul Rosen

paul.rosen@utah.edu @paulrosenphd https://cspaul.com

Visualization for Data Science DS-4630 / CS-5630 / CS-5630 / CS-6630

Visualization for Data Science Tasks, Design and Evaluation



THE UNIVERSITY OF UTAH

Tasks Analysis



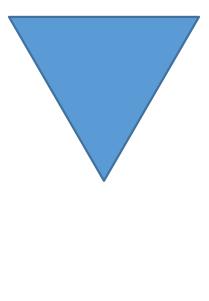


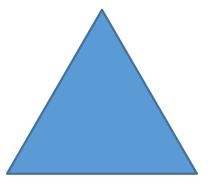
U

Problem-Driven vs Technique- Driven

- problem-driven
 - top-down approach
 - identify a problem encountered by users
 - design a solution to help users work more effectively sometimes called a design study
- technique-driven
 - bottom-up approach
 - invent new visualization techniques or algorithms
 - classify or compare against other idioms and algorithms











IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 15, NO. 6, NOVEMBER/DECEMBER 2009

A Nested Model for Visualization Design and Validation

Tamara Munzner, Member, IEEE

Abstract—We present a nested model for the visualization design and validation with four layers: characterize the task and data in the vocabulary of the problem domain, abstract into operations and data types, design visual encoding and interaction techniques, and create algorithms to execute techniques efficiently. The output from a level above is input to the level below, bringing attention to the design challenge that an upstream error inevitably cascades to all downstream levels. This model provides prescriptive guidance for determining appropriate evaluation approaches by identifying threats to validity unique to each level. We also provide three recommendations motivated by this model: authors should distinguish between these levels when claiming contributions at more than one of them, authors should explicitly state upstream assumptions at levels above the focus of a paper, and visualization venues should accept more papers on domain characterization.

Index Terms—Models, frameworks, design, evaluation.

1 INTRODUCTION

Many visualization models have been proposed to guide the creation and analysis of visualization systems [8, 7, 10], but they have not been tightly coupled to the question of how to evaluate these systems. Similarly, there has been significant previous work on evaluating visualization [9, 33, 42]. However, most of it is structured as an enumeration of methods with focus on *how* to carry them out, without prescriptive advice for *when* to choose between them.

The impetus for this work was dissatisfaction with a flat list of evaluation methodologies in a recent paper on the process of writing visualization papers [29]. Although that previous work provides some guidance for when to use which methods, it does not provide a full framework to guide the decision or analysis process.

In this paper, we present a model that splits visualization design into levels, with distinct evaluation methodologies suggested at each level based on the threats to validity that occur at that level. The four levels are: characterize the tasks and data in the vocabulary of the problem domain, abstract into operations and data types, design visual encoding and interaction techniques, and create algorithms to execute these techniques efficiently. We conjecture that many past visualization designers did carry out these steps, albeit implicitly or subconsciously, and not necessarily in that order. Our goal in making these steps more systems, and compare our model to previous ones. We provide recommendations motivated by this model, and conclude with a discussion of limitations and future work.

2 NESTED MODEL

Figure 1 shows the nested four-level model for visualization design and evaluation. The top level is to characterize the problems and data of a particular domain, the next level is to map those into abstract operations and data types, the third level is to design the visual encoding and interaction to support those operations, and the innermost fourth level is to create an algorithm to carry out that design automatically and efficiently. The three inner levels are all instances of design problems, although it is a different problem at each level.

These levels are nested; the output from an *upstream* level above is input to the *downstream* level below, as indicated by the arrows in Figure 1. The challenge of this nesting is that an upstream error inevitably cascades to all downstream levels. If a poor choice was made in the abstraction stage, then even perfect visual encoding and algorithm design will not create a visualization system that solves the intended problem.





Purpose of the Nested Model

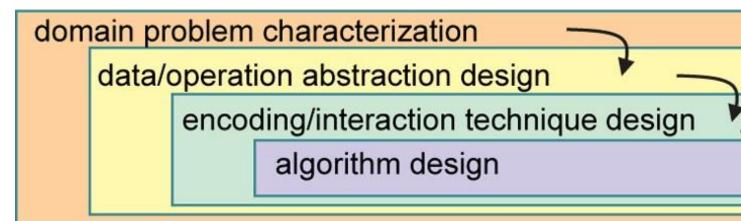
- capture design decisions
 - what is the justification behind your design?
- analyze aspects of the design process
 - broken apart into four different concerns
- validate early & often
 - avoid making ineffective solutions





U

Nested Model for Visualization Design



threat: wrong problem validate: observe and interview target users threat: bad data/operation abstraction threat: ineffective encoding/interaction technique validate: justify encoding/interaction design threat: slow algorithm validate: analyze computational complexity implement system validate: measure system time/memory validate: qualitative/quantitative result image analysis [test on any users, informal usability study] validate: lab study, measure human time/errors for operation validate: test on target users, collect anecdotal evidence of utility validate: field study, document human usage of deployed system validate: observe adoption rates





Design

Threats & Evaluation

Munzner 2009

Design Process

Understand Map to Identify & Implement Abstract Task Suitable Technique

Data Type & Other Factors

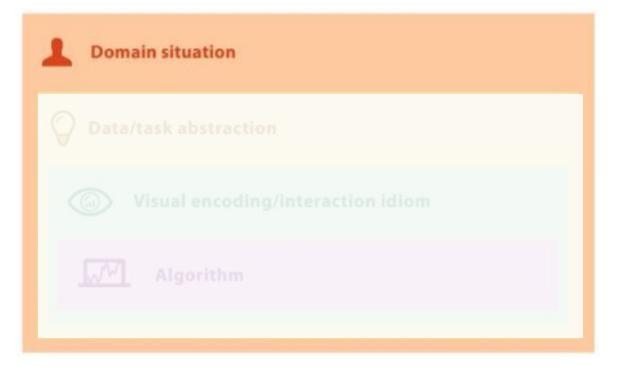




Domain Characterization

- details of an application domain
- group of users, target domain, their questions, & their data
 - varies wildly by domain
 - must be specific enough to continue with
- cannot just ask people what they do
 - introspection is hard!





Domain Problem Characterization

- Infinite numbers of domain tasks
- Can be broken down into simpler abstract tasks
- We know how to address the abstract tasks!
- Identify task data combination: solutions probably exist





Example: Find Good Movies

- I want to identify good movies in genres I like.
- Domain: general population, movie enthusiasts



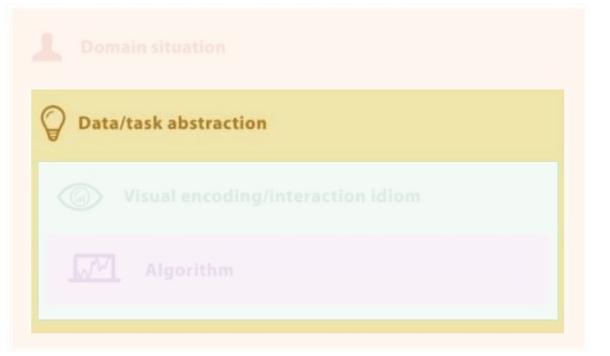


U

Data & Task Abstraction

- the what-why, map into generalized terms
- identify tasks that users wish to perform or already do
- find data types and good model of the data
- sometimes must transform the data for a better solution
 - this can be varied and guided by the specific task





Example: Find Good Movies

- What is a good movie for me?
 - Highly rated by critics?
 - Highly rated by audiences?
 - Successful at the box office?
 - Similar to movies I liked?
 - Specific Genres?
- Data Sources: IMDB, Rotten Tomatoes, ...



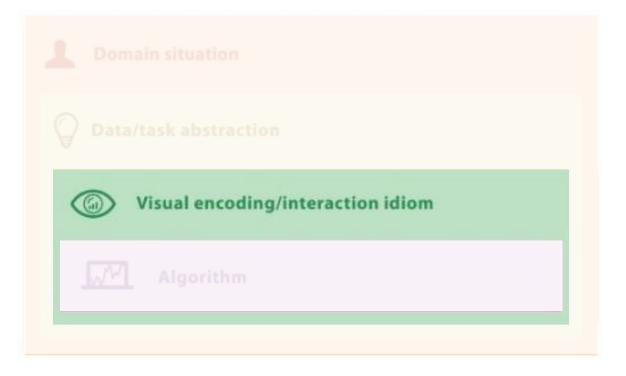


U

Encodings & Interactions

- the design of visualization techniques
 - visual encodings interactions
- ways to create and manipulate the visual representation of data
- decisions on these may be separate or intertwined
- visualization design principles drive decisions





Example: Find Good Movies

- Combination of audience ratings and critics ratings, filtered by genre.
- Idiom: stacked bar chart for ratings filter interface for genre



Rank	T Institution	⊛ national_rank (2012) (50) 1. [3] ☴	T Country
	+z ~ ·	# national_ran # guality_of_e	+Z ~~ (
		μοτημοτική μοτημα ματική μοτημα ματική μοτημα ματικά μα μι μα μοτημα ματικά	2
10		1.00 58.001.00 101.00	
10	University of Texas MD Anderson		USA
20	Georgia Institute of Technology		USA
	University of Virginia		USA
	Emory University		USA
	Arizona State University		USA
	Texas A&M University, College S		USA
	Rutgers University-New Brunswig		USA
	Pennsylvania State University, Ur		USA
	Boston University		USA
	University of Florida		USA
	Ohio State University, Columbus		USA
	University of Maryland, College F		USA
	University of Arizona		USA
	University of Pittsburgh - Pittsbu		USA
	University of Minnesota, Twin Cit		USA
	University of California, Irvine		USA
	University of Southern California		USA
18 🗆	Tufts University		USA
	Vanderbilt University		USA
20 🗆	University of California, Santa Ba		USA
21 🗆	University of Washington - Seattl		USA
22 🗆	Purdue University, West Lafayett		USA
23 🗆	University of North Carolina at C		USA
24 🗆	University of Colorado Boulder		USA
25 🗆	Northwestern University		USA
	University of Texas at Austin		USA
	University of Utah		USA
	University of California, Davis		USA
	University of California, San Fran		USA
	Brown University		USA
	Case Western Reserve University		USA
	Washington University in St. Lou		USA
	Williams College		USA
	New York University		USA
	University of Illinois at Urbana-C		USA
	University of Bristol		United Kingdom
	University of Nottingham		United Kingdom

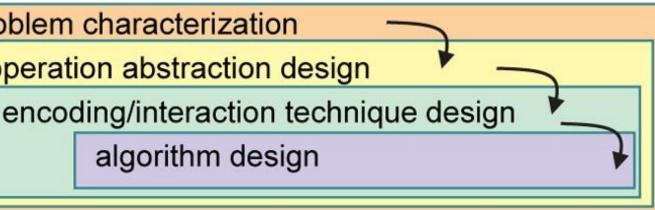
Task Abstraction Exercise

You have been approached by a geneticist to help with a visualization problem. She has gene expression data (data that measures the activity of the genes) for **30 cancer tissue samples**. She is applying an experimental drug to see whether the cancer tissue dies as she hopes, but she finds that only some samples show the desired effect. She believes that the difference between the samples is caused by differential expression (different activity) of genes in a particular pathway, i.e., an interaction network of genes. She would like to understand which genes are likely to cause the difference, and what role they play in that pathway.

- Objective 1: Task Abstraction
- Objective 2: Encoding Design

domain problem characterization data/operation abstraction design





Task Abstraction

...only some samples show the desired effect.
 -> derive two groups of samples





Task Abstraction

• ... the difference between the samples is caused by differential expression (different activity) of genes in a particular pathway. She would like to understand which genes are likely to cause the difference

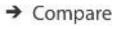
-> identify those genes

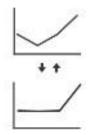
- -> compare gene expression of pathway genes between two groups
- -> identify the outliers

→ Identify









Outliers





Task Abstraction

• She would like to understand which genes are likely to cause the difference, and what role they play in that pathway.

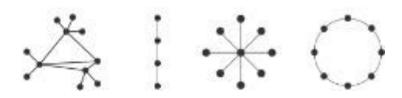
-> Locate the outlier in the network

-> Explore the topology



	Target known	
Location known	•_•• Lookup	
Location unknown	COCOL Locate	

→ Topology





Encoding Design

• Heat Maps?

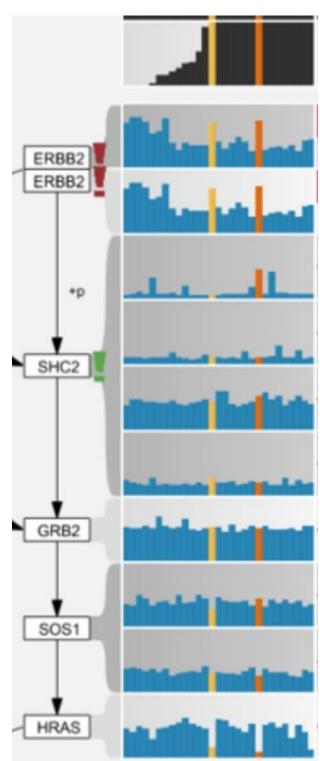
Bar Charts?

- Tabular Data, 30 samples, 30 genes Compare groups, spot outliers
 - Dimensionality Reduction?
 - Scatterplot Matrices?
 - Parallel Coordinates?

- compare groups. 30 Dimensions is too much -> Scalability comparison necessary
- Work! Spatial separation of groups.
- Work even better! 30x30 still feasible, encoding advantage



- Doesn't show raw data, not great to
- 30 Dimensions is a lot, coloring for



Encoding Design

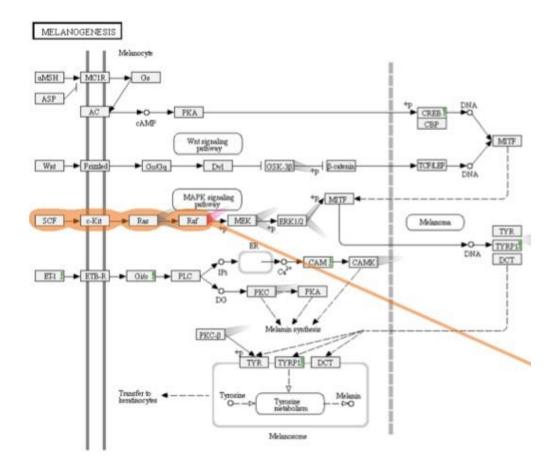
- Network, 30 genes
- Explore Topology, Lookup Nodes
 - Matrix?
 - Treemap?

• Node-Link Diagram?

Doesn't work for topology tasks

Works well.





- Doesn't work for general networks
- Combine with Table through highlighting.



Designing Visualizations





What is Design?

- creating something new to solve a problem
- can be used to make buildings, chairs, user interfaces, etc.
- design is used in many fields
- many possible users or tasks





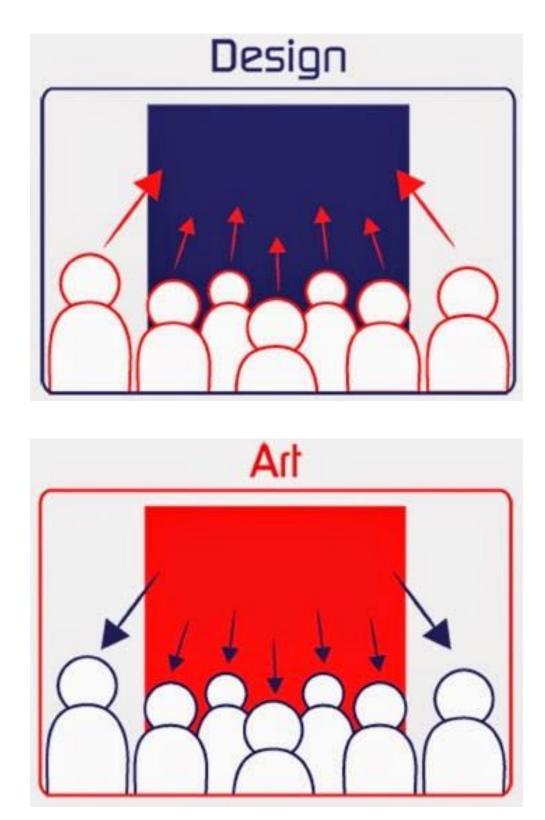
https://www.youtube.com/watch?v=hUhisi2FBuw



What is Design Not?

- just making things pretty
- art appreciation of beauty or emotions invoked
- something without a clear purpose
- building without justification or evidence





http://woodyart211.blogspot.com/2015/01/art-vs-design-comments.html



Form & Function

- commonly: "form follows function"
- function can constrain possible forms
- form depends on tasks that must be achieved
- "the better defined the goals of an artifact, the narrower the variety of forms it can adopt"



http://img.weburban main-960x481.jpg

-Alberto Cairo



The Functional Art: An introduction to information graphics and visualization. New Riders, 2012.

http://img.weburbanist.com/wp-content/uploads/2015/05/sculptural-furniture-



When do we Design?

- wicked problems
 - no clear problem definition
 - solutions are either good enough or not good enough
 - multiple solutions exist, not true/false no clear point to stop with a solution
- examples of non-wicked ("tame") problems
 - mathematics, chess, puzzles





Tacoma Narrows Bridge

Dilemmas in a general theory of planning. Rittel, H.W. and Webber, M.M., Policy Sciences, 1973.

Why does Design Matter for Vis?

- many ineffective visualization combinations
- users with unique problems & data
- variations of tasks
- large design space



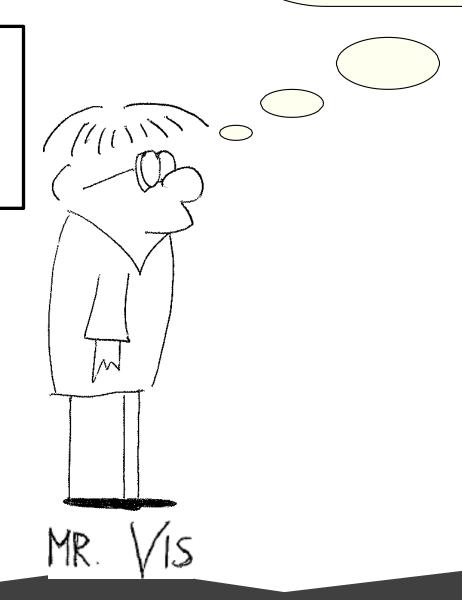
?

U



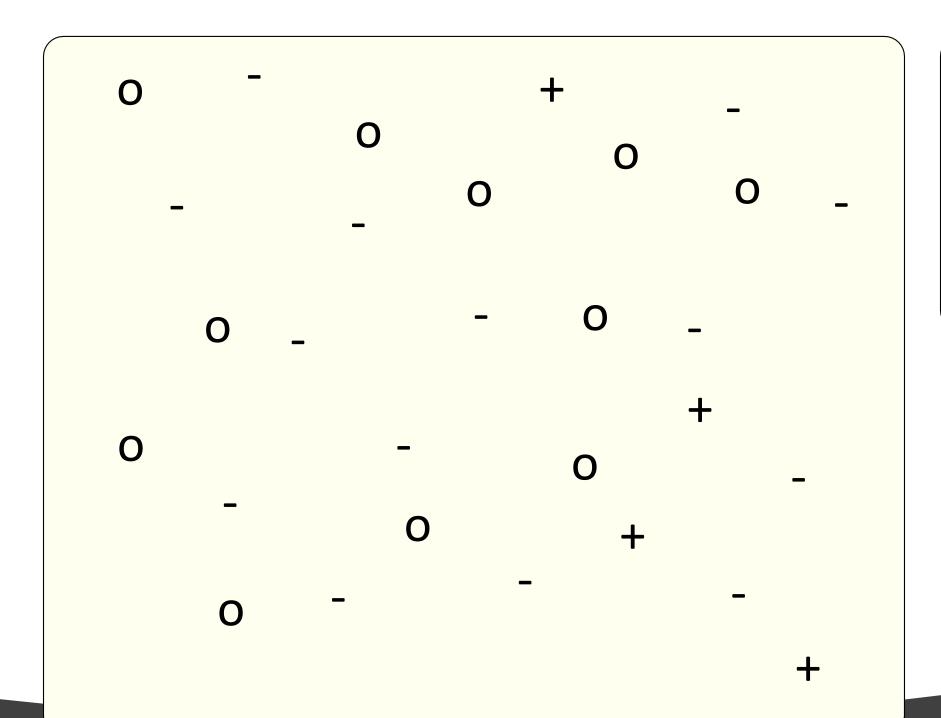
Of course they need the cool **technique** I built last year!

PREMATURE DESIGN COMMITMENT





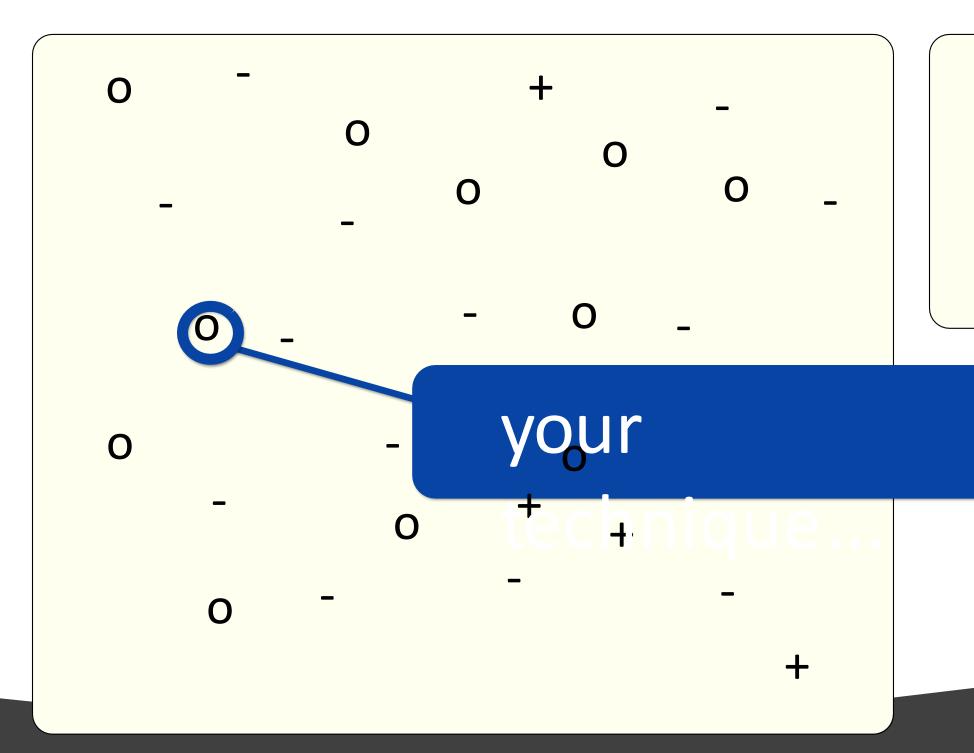






good okay poor

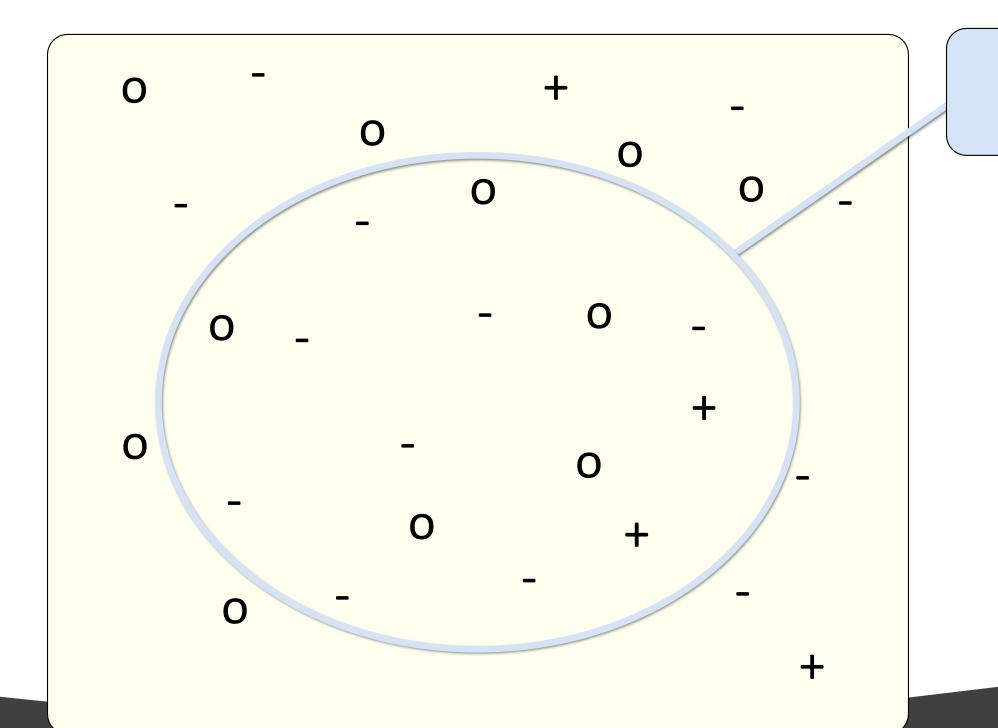






good okay poor



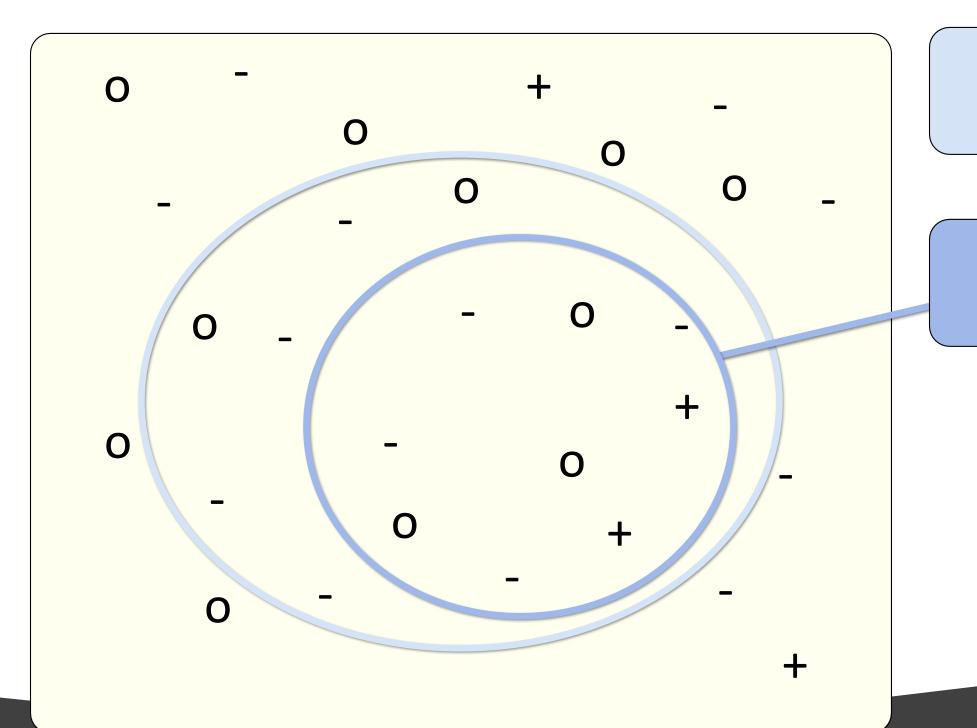




kno

W

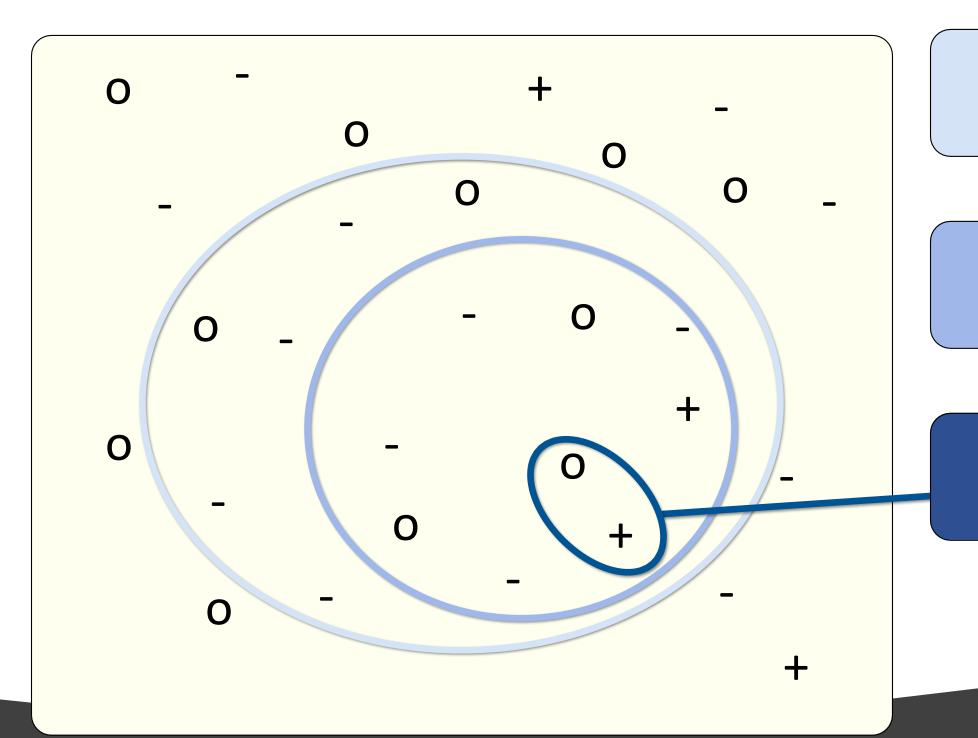






kno w conside r

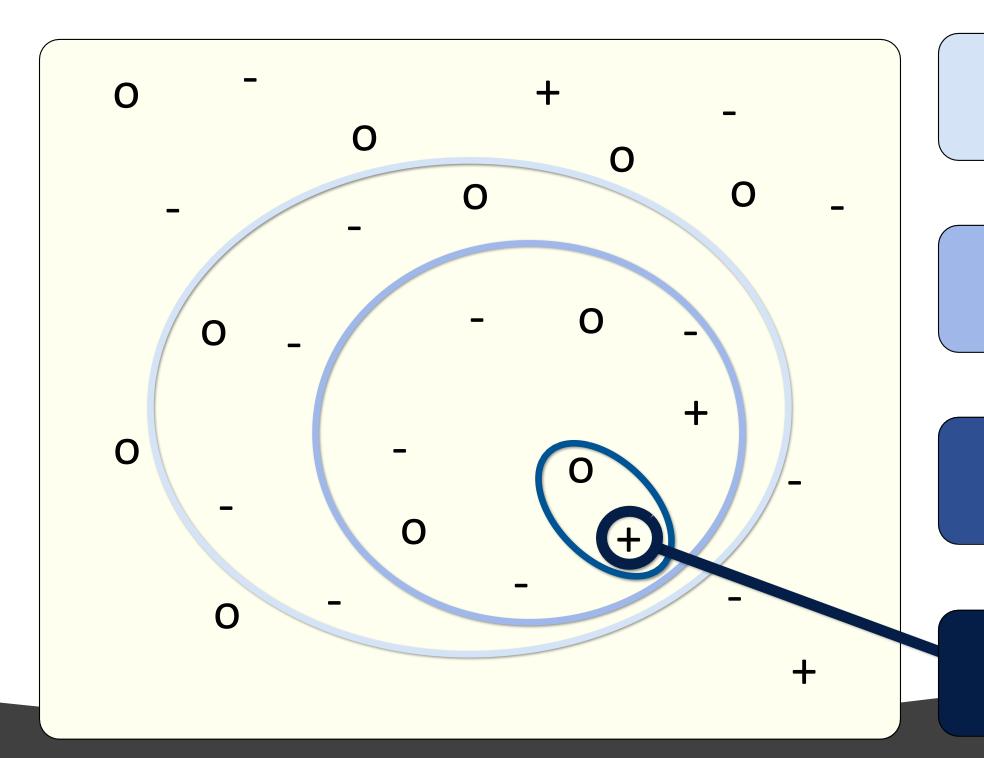






kno w conside r propos







kno w conside r propos

selec

+







+ good o okay - poor

conside

propose



Design Methods

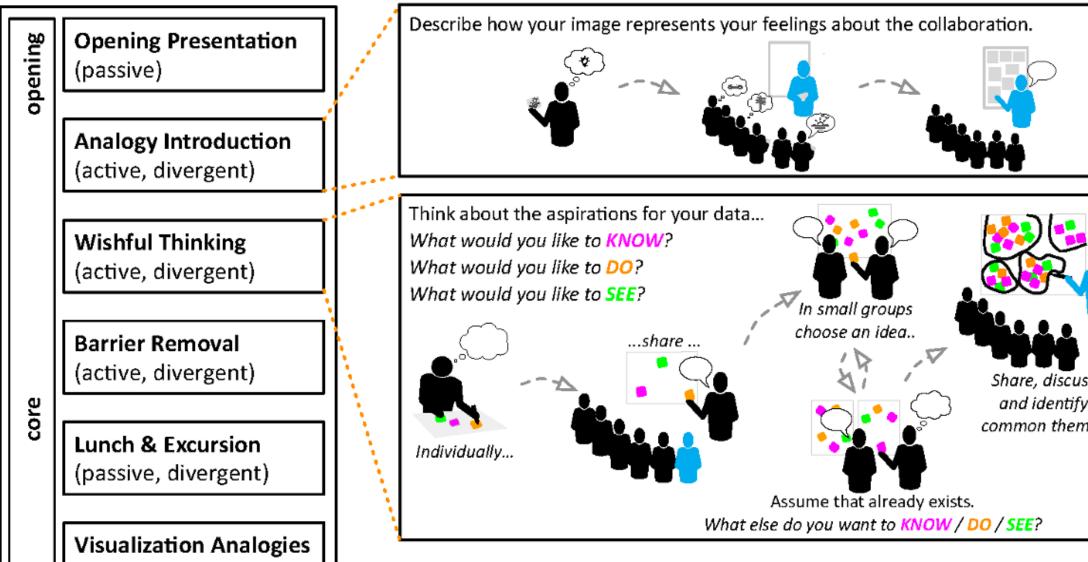




U

Creativity Workshops

- goals:
 - generate design requirements
 - promote creativity
- combined a variety of techniques:
 - wishful thinking
 - constraint removal
 - excursion
 - analogical reasoning
 - storyboarding
- measured prototypes for appropriateness, novelty, & surprise





Ethan Kerzner, Sarah Goodwin, Jason Dykes, Sara Jones, Miriah Meyer A Framework for Creative Visualization-Opportunities Workshops IEEE Transactions on Visualization and Computer Graphics (InfoVis '18) 2018

http://vdl.sci.utah.edu/CVOWorkshops/





Parallel Prototyping

- Develop multiple designs in parallel
- Example: graphic design
- serial vs parallel design: create & critique



serial

parallel





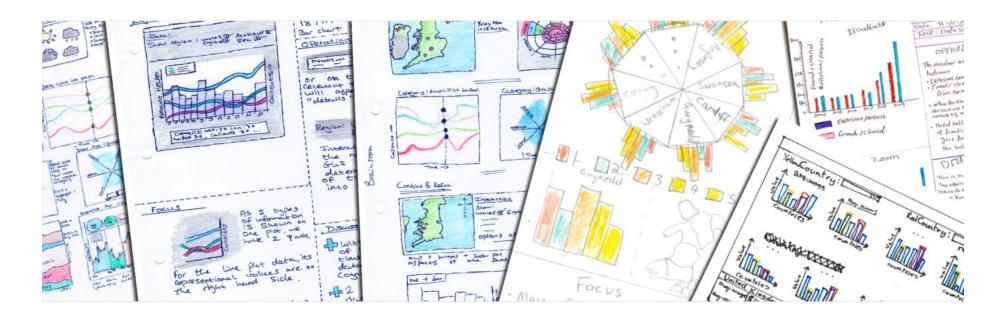
Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. Dow, S.P., Glassco, A., Kass, J., Schwarz, M.,

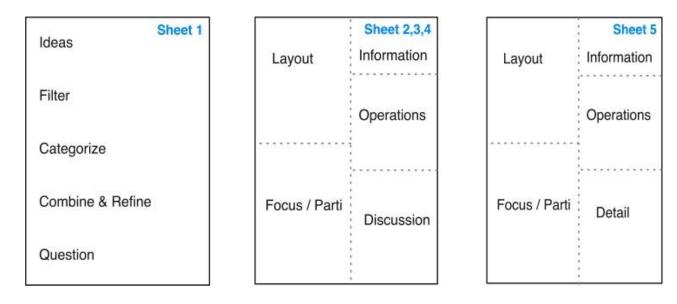
Five-Design Sheets

- tailored to visualization design
 - in industry and classroom use sketching as a way to plan
- the design sheets:
 - #1 brainstorm solutions to a task #2-4 different principle designs
 - #5 converge on design to implement

http://fds.design/



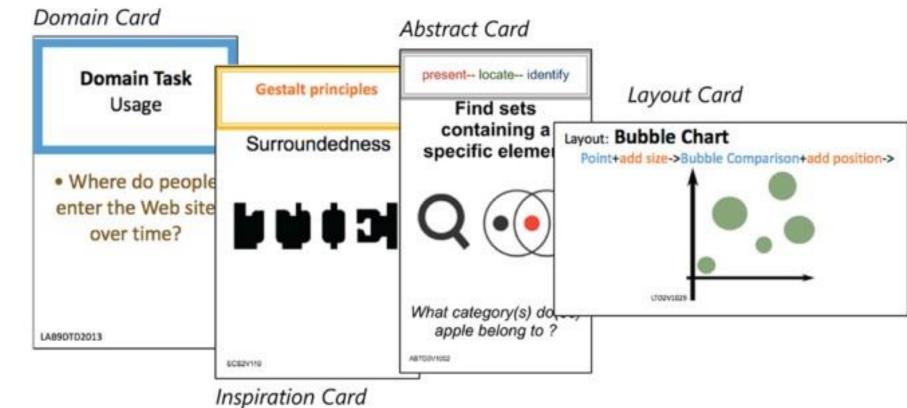




Sketching designs using the Five Design-Sheet methodology. Roberts, J.C., Headleand, C. and Ritsos, P.D., IEEE InfoVis, 2015.

Vizlt Cards

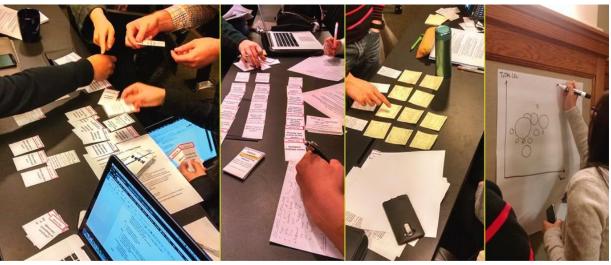
- different cards to assist with visualization design
- types of cards
 - domain
 - inspiration
 - abstract
 - layout



• aim to help students design, compare, collaborate, apply, and synthesize

http://vizitcards.org





VizIt Cards: A card-based toolkit for infovis design education. He, S. and Adar, E., IEEE InfoVis, 2016.

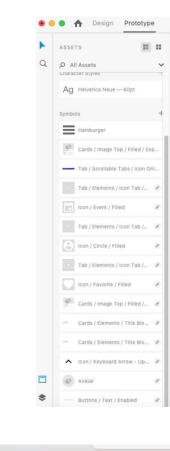


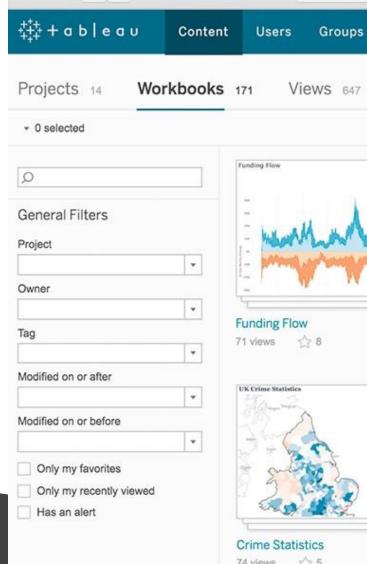
Wireframing

• Dedicated Tools like Figma, Adobe XD, or Sketch

• PowerPoint, Keynote, Illustrator

Need Data: Tableau



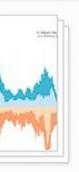




Untitled-1	-	Edi	ted

	Untitled-1 — Ed	aited	32.1% 🖌 📘 🕨 🛄
Web 1920 - 1		iPhone X – 1	iPhone X - 2
		WELCOME TO OUR WEBSITE	EEPINOE TITLE
		GET STARTED	
	WELCOME TO OUR WEBSITE		A quice description of the service of a manufacture of the service cost guide the table from a
	GET STARTED	OVA BERVICES	
			Jan makazo Lan or asasing
		SERVICE TITLE	- Nori-Aufler 1 - Rein-Aufler 1 - Rein-Aufler 1 - Rein-Aufler 2 - Rein-
	•		
	-		
	OUR SERVICES		
_			
	- Nutlina 1		
Headine 6 Headine 6 House 2 Known and developments of magnetic for.	Meadine 6 Headine 6	Edit Rotate Mask Scale Fielder	Union Subtract Intersect Difference Zoom View Pr
arrest portes		(100 CONT	74.5 ×
	Document		226 W 🛱
	iOS UI Design	₽	PROTOTYPING
	V Components	CATES	SYMBOL
	About		Regular Document
	E Card		22 Detach
			Overrides
-			Label 👳 🦂
Schedules	Tasks Status Se	ttings O	APPEARANCE
			Opecity (Norma
			STYLE
			STYLE Shadows

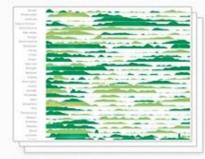
Data Sources 32





61 views 🖧 3

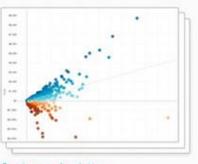




Unemployment

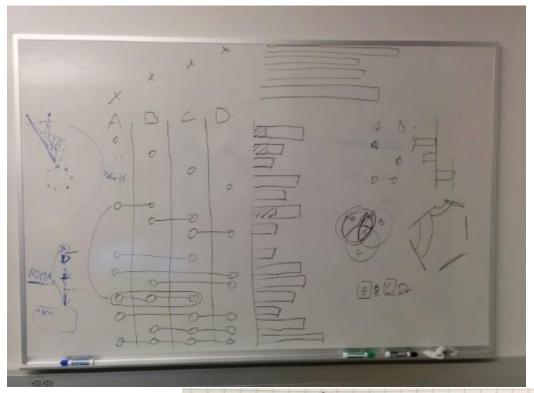


Census 2013 Median Age Views 395 views 🛛 🐈 15



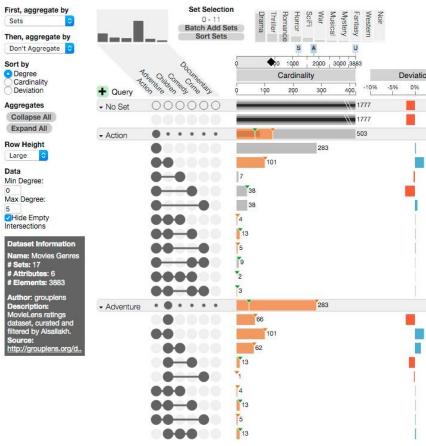
Customer Analytics 4 284 vioue





	E472 PO4304 BAA7 BAA7 RA2		
	F 2000	COUNT EXPE	CT DEVIL
EGR	1-557 V	and the	2r E
PDGTRA	0		P
RAS ==	•	= =	
BRAJ=	0		
BRCAT=	0		E
RAL >	0	the Bern	- El-
OR7 >	20SET V	122028 22281	- FE
SRI =	0# 00		12-
	0-0		
	0-0		D
	0-0	1	4
			-
		- A	4
	0	2 3	P
	0-0		E
	3-SET A	122201 220	
	4-SET A	enen: 31	F
	5-SET A	3379 7	23
	DETAILS		
	P.0 0 2 #	PATIENT	1
	0 0 0	PATIENT	
	10	PATKENT	
	LOGIC		
	0 0 70		E
	The the		2

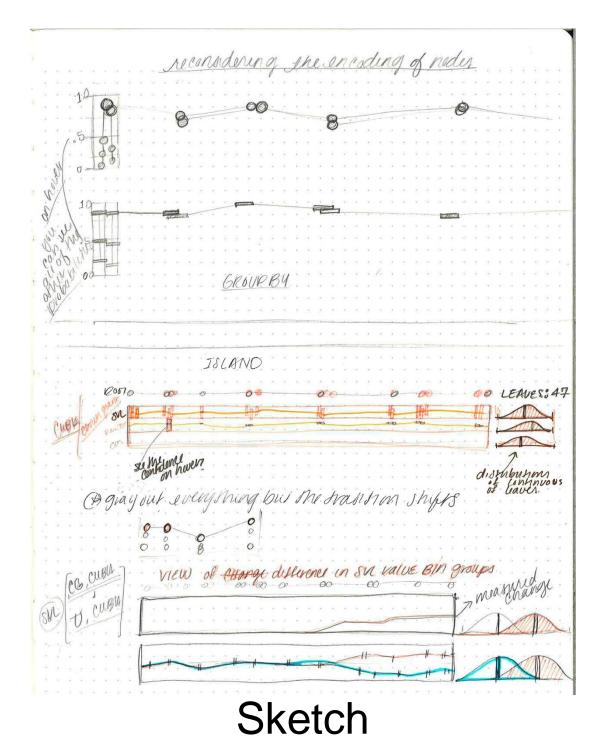
UpSet - Visualizing Intersecting Sets

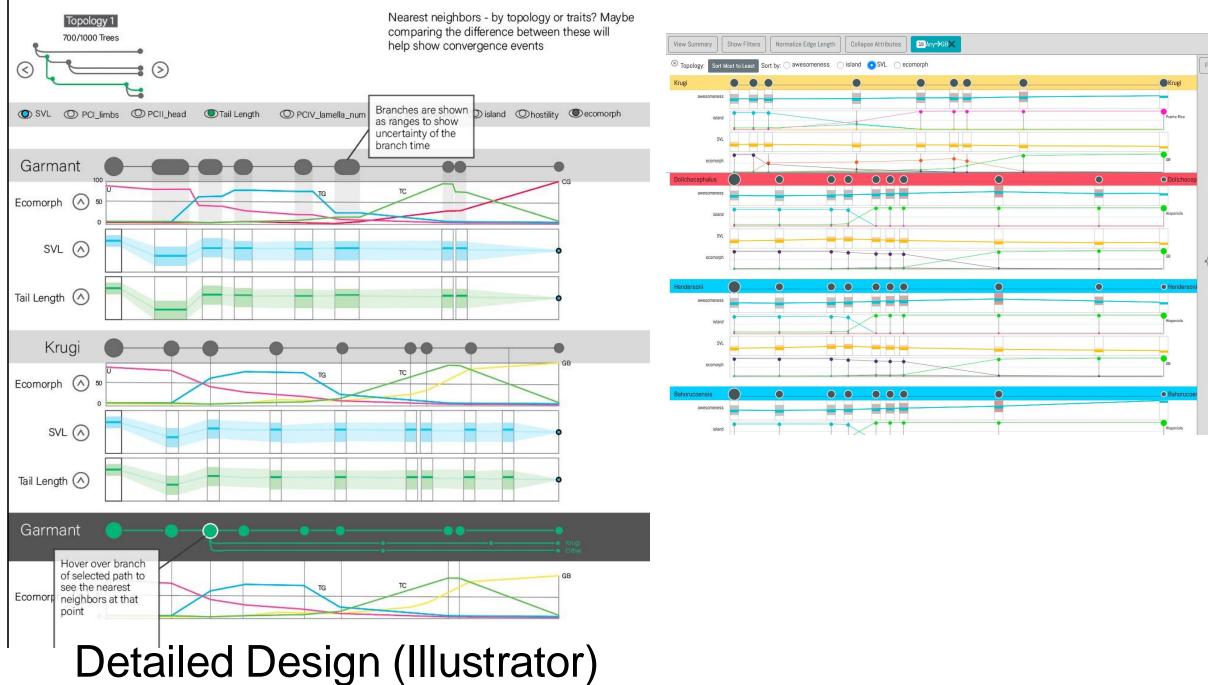




	Choose Datas	Set Movies Genres	(17 sets, 4 attributes)			out UpS	et UpSet	t for F
tion	Release Date	Average Rating	Times Watchec	Element Visua 5.0 4.5 8.0 4.5 8.0 4.0 3.5 3.0 2.5 -				
	H			2.0 -			1.1	$\mathcal{A}_{\mathcal{A}}$
				1.0				
				1,920 1,930 Release Dat		0 1,960 1,9	70 1,980 1,9	90 2,00
				+ Scatterplo	•		()	-
	Hite Hite			B				
	יער ורח–ו		₩ T	Element Quer				
					and Commencement	5		
	. ц.		Н	<mark> </mark> 283 🕲 🗧	1200 3			
	ИШН	⊢-œ	нш—н	+				
	н)и	на	⊢₽	Query Filters				
	НĴł	НШН		O & Subset	L Sets			
	H	н	Н	000000	10010			
	н	H-I	Н	+ Name		Contains		
	H							
				Query Results	CACHONIC	12		
				Name	Release Date	Average Rating	Times Watched	Set Cour
				Jumanji (1995)	1995	3.2	701	3
	I H		I Н	Tom and Huck (1995)	1995	3.01	68	2
	ЮН	⊢-ф	⊢∰⊒I ⊢∰I	GoldenEye (1995)	1995	3.54	888	3
	HJ H <u>I</u> III		⊢∎−−1 ₽₽−−1	Cutthroat Island (1995)	1995	2.46	146	3

Implementation





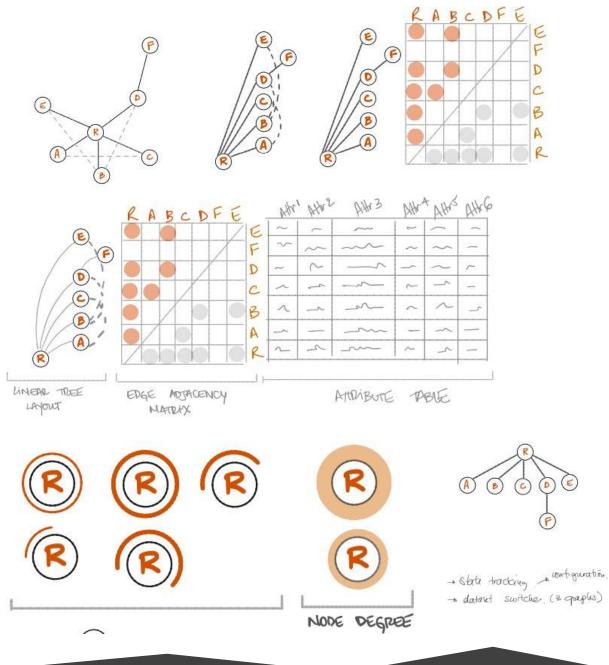


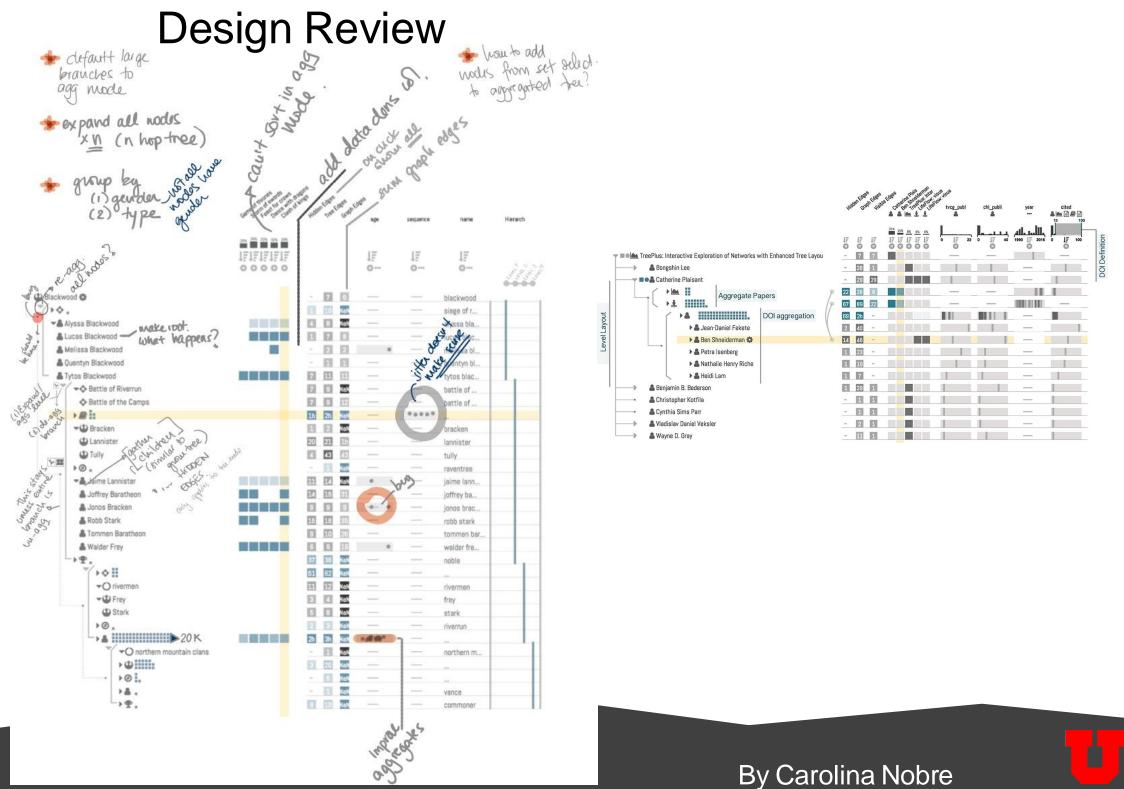
By Jen Rogers



Implementation

Sketch







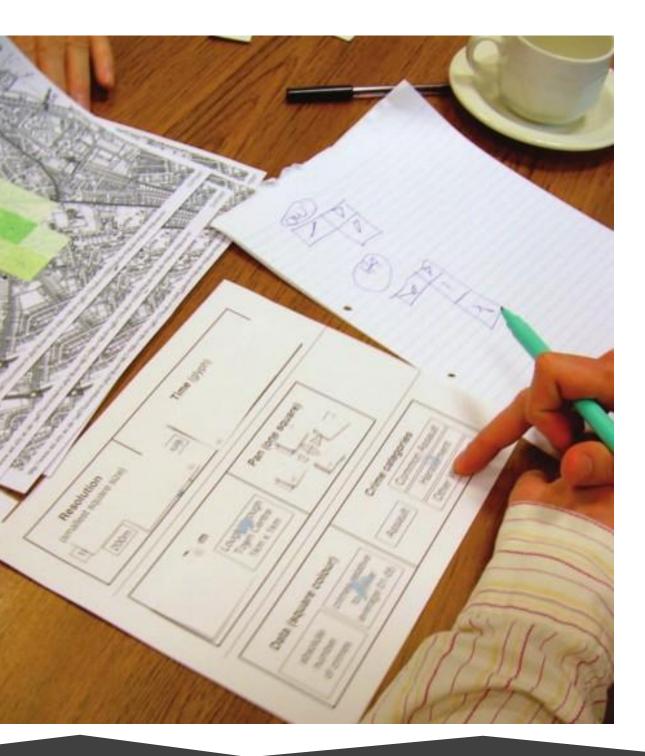
Interactive Prototyping

- "create a paper-based simulation of an interface to test interaction with a USER"
 Methods to support human-centred design. Maguire, M., International Journal of Human-Computer Studies, 2001
- received more suggestions than digital

users requested more features to add

 hypothesis that paper prototyping stimulates creativity and interaction





Human-centered approaches in geovisualization design: Investigating multiple methods through a long-term case study. Lloyd, D, and Dykes, L, IEEE Info//is, 2011



Other Methods

- interviews/observations
- qualitative analysis
- personas
- data sketches
- coding





U

Evaluation





U

Evaluating Information Visualizations

Sheelagh Carpendale

Department of Computer Science, University of Calgary, 2500 University Dr. NW, Calgary, AB, Canada T2N 1N4 sheelagh@ucalgary.ca

Introduction 1

Information visualization research is becoming more established, and as a result, it is becoming increasingly important that research in this field is validated. With the general increase in information visualization research there has also been an increase, albeit disproportionately small, in the amount of empirical work directly focused on information visualization. The purpose of this paper is to increase awareness of empirical research in general, of its relationship to information visualization in particular; to emphasize its importance; and to encourage thoughtful application of a greater variety of evaluative research methodologies in information visualization.

One reason that it may be important to discuss the evaluation of information visualization, in general, is that it has been suggested that current evaluations are not convincing enough to encourage widespread adoption of information visualization tools [57]. Reasons given include that information visualizations are often evaluated using small detected with university student participants and using simple tasks. To an





Role of Evaluation / Validation

- Goals:
 - avoid ineffective solutions justify solutions
- Unique vs Comparison
 - others?
 - Is mine one of a kind?

- Dimensions:
 - Perception vs Technique/System
 - Is size a better visual channel than angle?
 - Is my visualization system any good?
- Usability Testing: • Check for problems with system





• Can I easily compare my vis to

Example: Three Linking Techniques Perception / Comparison

Straight Visual Links



present building occupied 1776, is the oldest Baptist

IPS map of the W3R Route

The section

Roger Williams, and was one of the original Thirteen Colonies of the United States.^[10] Although the west bank of the Providence River was later claimed by both the English and the Dutch, the actual inhabitants and true masters of the region were the Pokanoket Tribe of the Wampanoag Nation led by Massasoit ^{1]} Williams secured permission to settle from the Pokanoket and gave the city its present name. [12] Williams' ice soon became a refuge for persecuted religious senters, as he himself had been exiled from Massachus ence's growth would be slow during the next quarter-century the subsuming of its territory into surrounding towns, difficulty of farming the land, and differing of local traditions and land conflicts

all slowed development In the mid-1770s, the British government levied taxes that impeded Providence's maritime, fishing and agricultural industries, the mainstay of the city's economy. One example was the Sugar Ac

which was a tax levied against Providence's distille hat adversely affected its trade in rum and sla These taxes caused Providence to join the other olonies in renouncing allegiance to the British response to enforcement of unpopular trade vidence residents shilled the first blood of the

occupation, the capture of nearby N pted industry and kept the population on alert. roops were quartered for various campaigns and wn University's University Hall was used as a parracks and military hospital.^[13]

After departing from Newport, French troops sent by King Louis XVI and commanded by the Comte de assed through Providence on their way to join the attack against British forces. The march idence was the beginning of a campaign led jointly by General George Wa a decisive march that ended with the defeat of General wallis in the Siege of Yorktown at Yorktown Virginia and the Battle of the Chesapeake.

Following the war. Providence was the country's i rgest city. b[)[13] with 7,614 people. The economy shifted from maritime endeavors to manufacturing, articularly machinery, tools, silverware, jewelry an extiles. By the turn of the twentieth century, dence boasted some of the largest manufa plants in the country including Brown & Shame

Google maps mexico Search Maps Send In Lin

Mexico - Google Maps

← → C ③ maps.google.com

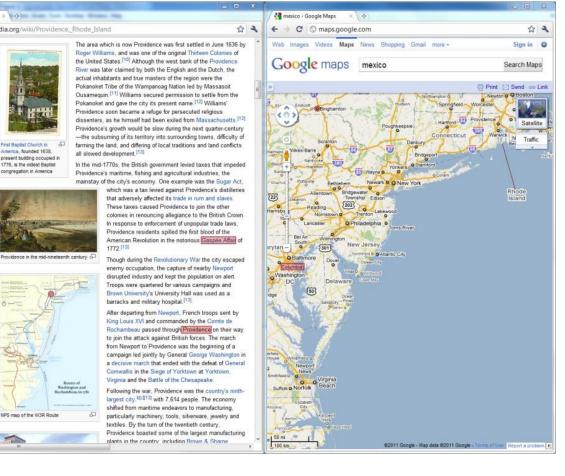
Web Images Videos Maps News Shopping Graail more

\$ 3



\$ 3

Sign in 🙀



Context-Preserving Visual Links



erica, founded 1638 1776, is the oldest Baptist





To BLACK

The area which is now Providence was first settled in June 1636 by r Williams, and was one of the original Ti the United States.^[10] Although the west bank of the Providence River was later claimed by both the English and the Dutch, the actual inhabitants and true masters of the region were the Pokanoket Tribe of the Wampanoag Nation led by Massasoi mequin.^[11] Williams secured permission to settle from the Pokanoket and gave the city its present name [12] Williams' dence soon became a refuge for persecuted religious dissenters, as he himself had been exiled from Massachuse ovidence's growth would be slow during the next quarter-century -the subsuming of its territory into surrounding towns, difficulty of farming the land, and differing of local traditions and land conflicts all slowed development [13]

In the mid-1770s, the British government levied taxes that imped Providence's maritime, fishing and agricultural industries, the mainstay of the city's economy. One example was the Sugar A

which was a tax levied against Providence's distille hat adversely affected its trade in rum and slave hese taxes caused Providence to join the other onies in renouncing allegiance to the British Cro response to enforcement of unpopular trade laws sidence residents shilled the first blond of the erican Revolution in the notorious Gaspée Affa

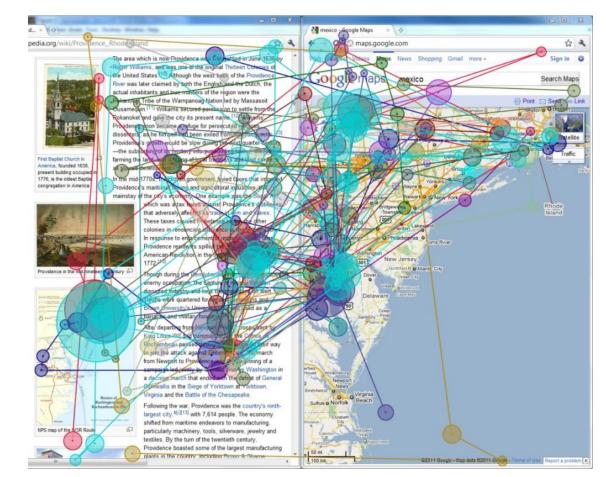
hough during the Revolutionary War the city es enemy occupation, the capture of nearby Newp disrupted industry and kept the population on alert roops were quartered for various campaigns and wn University's University Hall was used as a arracks and military hospital.[13]

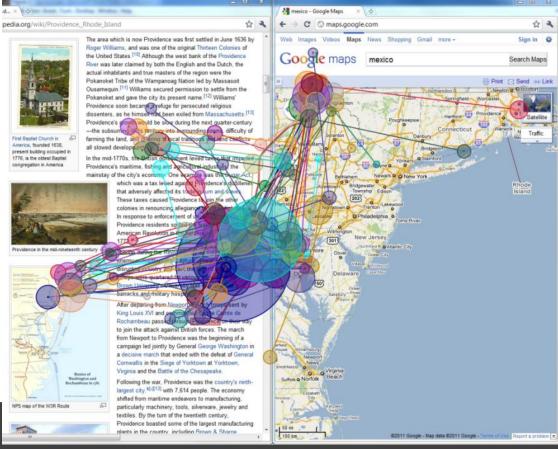
After departing from Newport, French tro King Louis XVI and commanded by the Camte de nbeau passed through Providence on their w to join the attack against British forces. The march from Newport to Providence was the beginning of a campaign led jointly by General George W a decisive march that ended with the defeat of Genera vallis in the Siege of Yorktown at Yorktown Virginia and the Battle of the Chesapeake

Following the war. Providence was the country's n argest city. b[/[13] with 7,614 people. The economy shifted from maritime endeavors to manufacturing inticularly machinery, tools, silverware, jewelry an extiles. By the turn of the twentieth century, rovidence boasted some of the largest manufact plants in the country including Brown & Shame



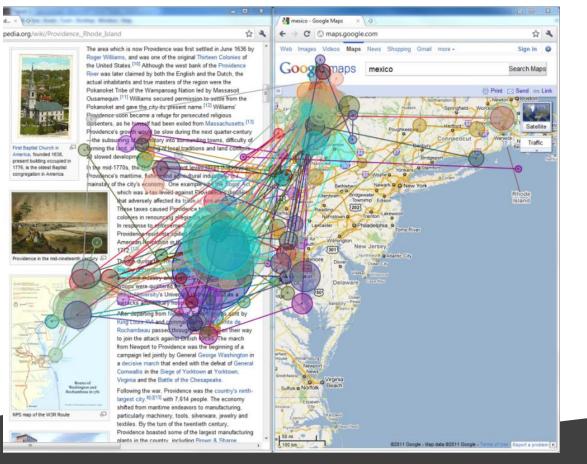
Gaze Plots Frame-Based Highlighting







Straight Visual Links





Context-Preserving Visual Links

What evaluation methods are there?

- Controlled experiment
 - Laboratory, Crowd-Sourced
- Interviews / questionnaires
 - Unstructured, structured, semi-structured
- Field observation, lab observation
 - Video / audio analysis
 - Coding / classification of user behavior (speech, gestures)
- Case studies



Certainty over what causes differences Not realistic (ecological validity)

High ecological validity Uncertain what causes differences



[Lam 2011]

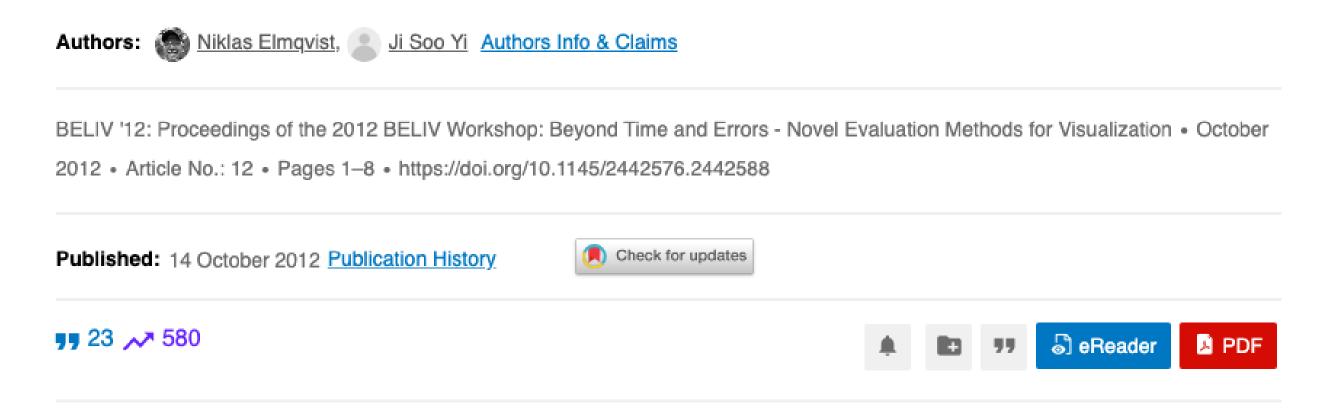
What evaluation methods are there?

- Algorithmic performance measurement
- Heuristic evaluation
 - Judge compliance with recognized metrics/usability methods (the heuristics)
- Usability testing, e.g., thinking aloud tests
- Wizard of Oz
 - Human simulates response of system
 - Test functionality before it's implemented
- Eye tracker evaluation
- Expert evaluation
- Insight-based evaluation
- Log analysis





Patterns for visualization evaluation



ABSTRACT

We propose a patterns-based approach to evaluating data visualization: a set of general and reusable solutions to commonly occurring problems in evaluating tools, techniques, and systems for visual sensemaking. Patterns have had significant impact in a wide array of disciplines, particularly software engineering, and we believe that they provide a powerful lens for looking at visualization evaluation by offering practical, tried-and-tested tips and tricks that can be adopted immediately. The 12 patterns presented here have also been added to a freely editable Wiki repository. The motivation for creating this evaluation pattern language is to (a) disseminate hard-won experience on visualization evaluation to researchers and practitioners alike: to (b) provide a standardized vocabulary for designing





Quantitative vs. Qualitative Evaluation

- Quantitative Methods
 - Objective metrics, measurements
 - Use numbers / statistics for interpreting data
- Qualitative Methods
 - Subjective metrics
 - Description of situations, events, people, interactions, and observed behaviors, the use of direct quotations from people about their experiences, attitudes, beliefs, and thoughts
 - Focused on understanding how people make meaning of and experience their environment or world





Typical Metrics

Objective Metrics

- Task completion time
- Errors (number, percent,...)
- Percent of task completed
- Ratio of successes to failures
- Number of repetitions
- Number of commands used
- Number of failed commands
- Physiological data (heart rate,...)
- Numbers of insights

Subjective Metrics

- Ratings
- Rankings
- User satisfaction
- Subjective performance
- Ease of use
- Intuitiveness
- Judgments

 \bullet

...

Comments and Feedback

•





Internal vs. External Validity

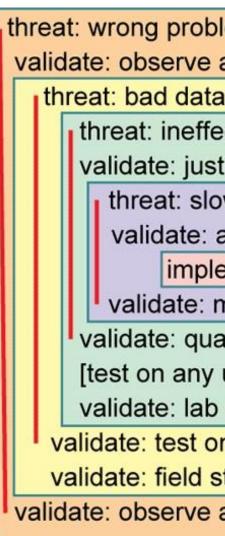
- Internal Validity can you trust your experiment
 - High when tested under controlled lab conditions
 - Observed effects are due to the test conditions (and not random variables)
- External Validity is your experiment representative of real world usage
 - High when interface is tested in the field, e.g. handheld device tested in museum
 - Results are valid in real world
- The Trade-off
 - The more akin to real-world situations, the more the experiment is susceptible to uncontrolled sources of variation





Scope of Evaluation

- Pre-design
 - e.g., to understand potential users' work environment and workflow
- Design
 - e.g., to scope a visual encoding and interaction design space based on human perception and cognition
- Prototype
 - e.g., to see if a visualization has achieved its design goals, to see how a prototype compares with the current state-of-the-art systems or techniques
- Deployment
 - e.g., to see how a visualization influences workflow and work processes, to assess the visualization's effectiveness and uses in the field
- Re-design
 - e.g., to improve a current design by identifying usability problems





lem		
and interview target users		
a/operation abstraction		
ective encoding/interaction technique	e	
tify encoding/interaction design		
ow algorithm		
analyze computational complexity		
ement system		
measure system time/memory		
alitative/quantitative result image and	alysis	
users, informal usability study]		
study, measure human time/errors for operation		
n target users, collect anecdotal evidence of utility		
study, document human usage of deployed system		
adoption rates		



Added value should be obvious!

Develop new methods/interface/software that are so awesome, cool, impressive, compelling, fascinating, and exciting that reviewers, colleagues, users are totally convinced just by looking at your work and some examples.

— Jarke van Wijk, Capstone Talk @ IEEE VIS 2013





More on this Topic

- CS 6540 HCI
- CS 6963 Advanced HCI
- ED PS 6010 Intro Statistics and **Research Design**
- DES 5710 Product Design and Development
- ANTH 6169 Ethnographic Methods
- ED PS 6030 Introduction to Research Design



In human-centered computing (HCC) the design and development of technology is motivated by the needs of people. HCC focuses on understanding how people use technology, creating new and accessible technology that enables novel interactions, and evaluating how technology impacts and supports people in the world. The core methods and techniques in HCC are grounded in computer science, but are also draw on social science and design. Current HCC focus areas in the School of Computing include personal informatics, mobile interaction, visualization, games, and privacy.

TRACK FACULTY

Michael Young

CORE CLASSES: Requir		
CS 6540	Н	
CS 6xxx	A	
CS 6630	V	
ED PS 6010	In	

ELECTIVES: 6 electives in total. tives require director approval

MS IN COMPUTING: HUMAN-CENTERED COMPUTING

Erik Brunvand, Rogelio E. Cardona-Rivera, Tamara Denning, Alexander Lex, Miriah Meyer (track director), Jason Wiese, R.

red courses:

dvanced HCI

isualization for Data Science

ntroduction to Statistics and Research Design

Pre-approved course list from within CS and across campus (1) Up to 3 electives can be taken from outside CS (2) Other elec-





U