

Methods for Visualizing Biodiversity & Building Rewarding Collaborations

Tamara Munzner

Department of Computer Science
University of British Columbia

UBC Biodiversity Challenge Retreat, Hakai Institute
11 June 2019

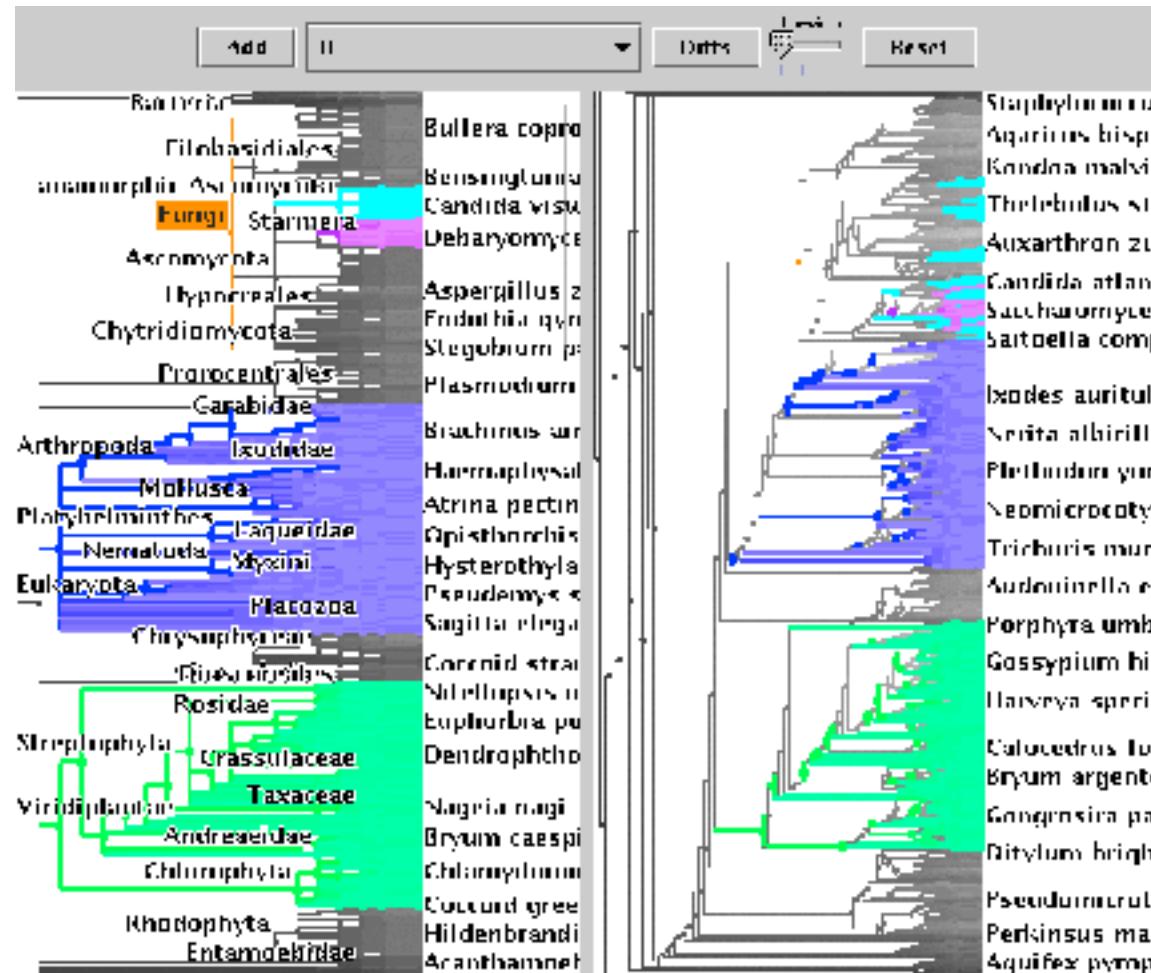
www.cs.ubc.ca/~tmm/talks.html#hakai19-methods



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

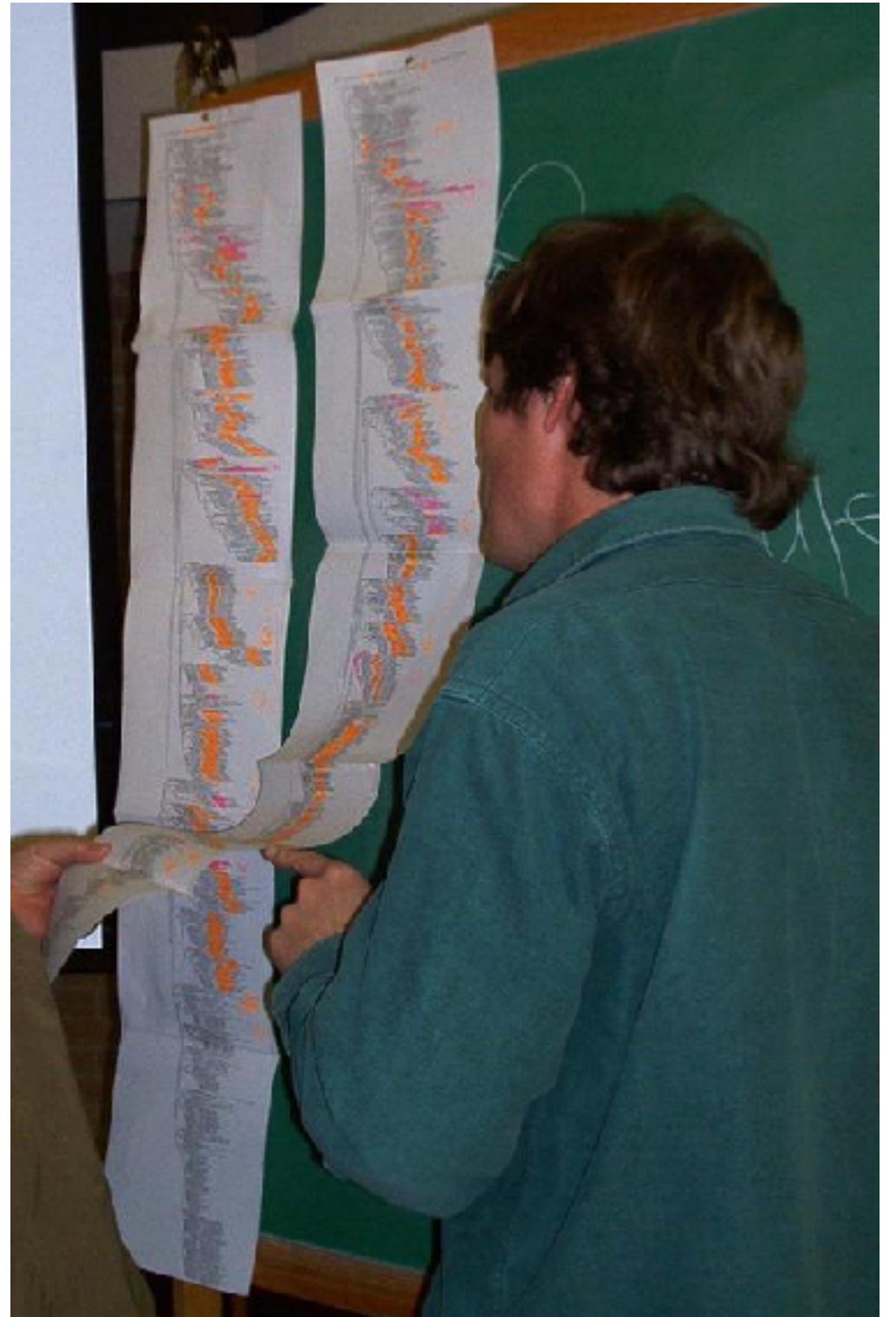
TreeJuxtaposer: Visual tree comparison

- driving problem from UT Austin Hillis Lab in 2001: phylogenetic trees
- algorithm focus on scale, later extended to gene sequences



Ruminant	C	G	N	N	N	N	N	N
Pig	C	A	CTACATC					
Llama	G	G	CTATATC					
Rhino	A	G	GTACATC					
Tapir	G	A	GTACATC					
Horse	G	A	GTACATC					
Cat	A	A	ATACATC					
Caniform	A	T	ATACATC					
Pangolin	G	G	ATACATC					

SequenceJuxtaposer

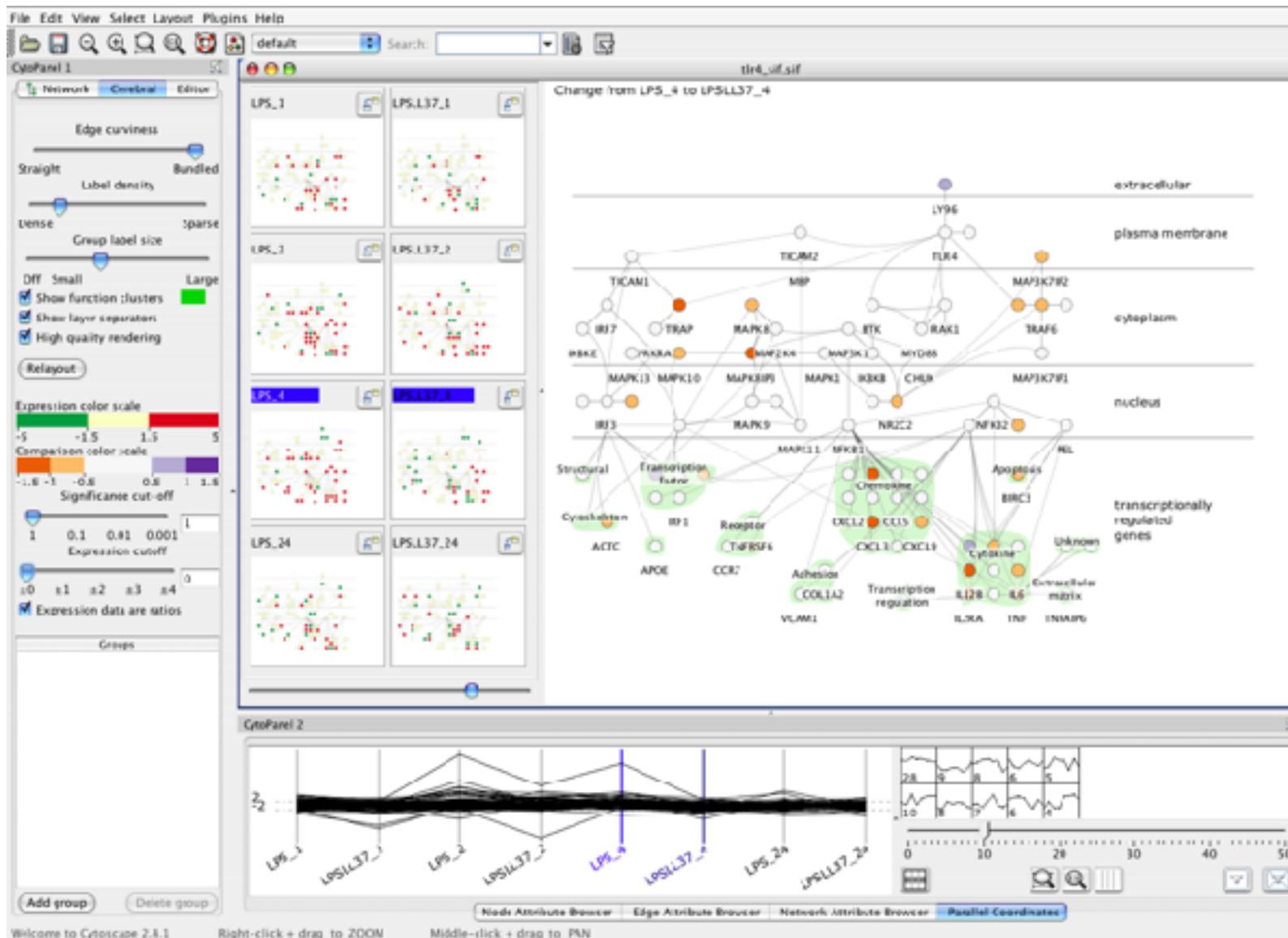


TreeJuxtaposer <https://youtu.be/GdaPj8a9QEo>

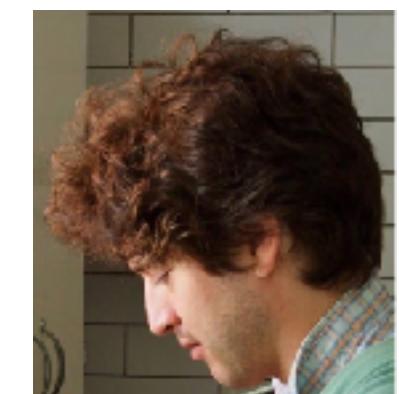
joint work with: Guimbretiere, Li, Zhang, and Zhou

Cerebral: Integrating gene expression w/ interaction network

- automatic network layout by subcellular location, like hand-drawn diagrams
- multiple views with linked highlighting and navigation
- Cytoscape plugin, funded by Agilent



Aaron Barsky



Jenn Gardy
(Microbio: Hancock)

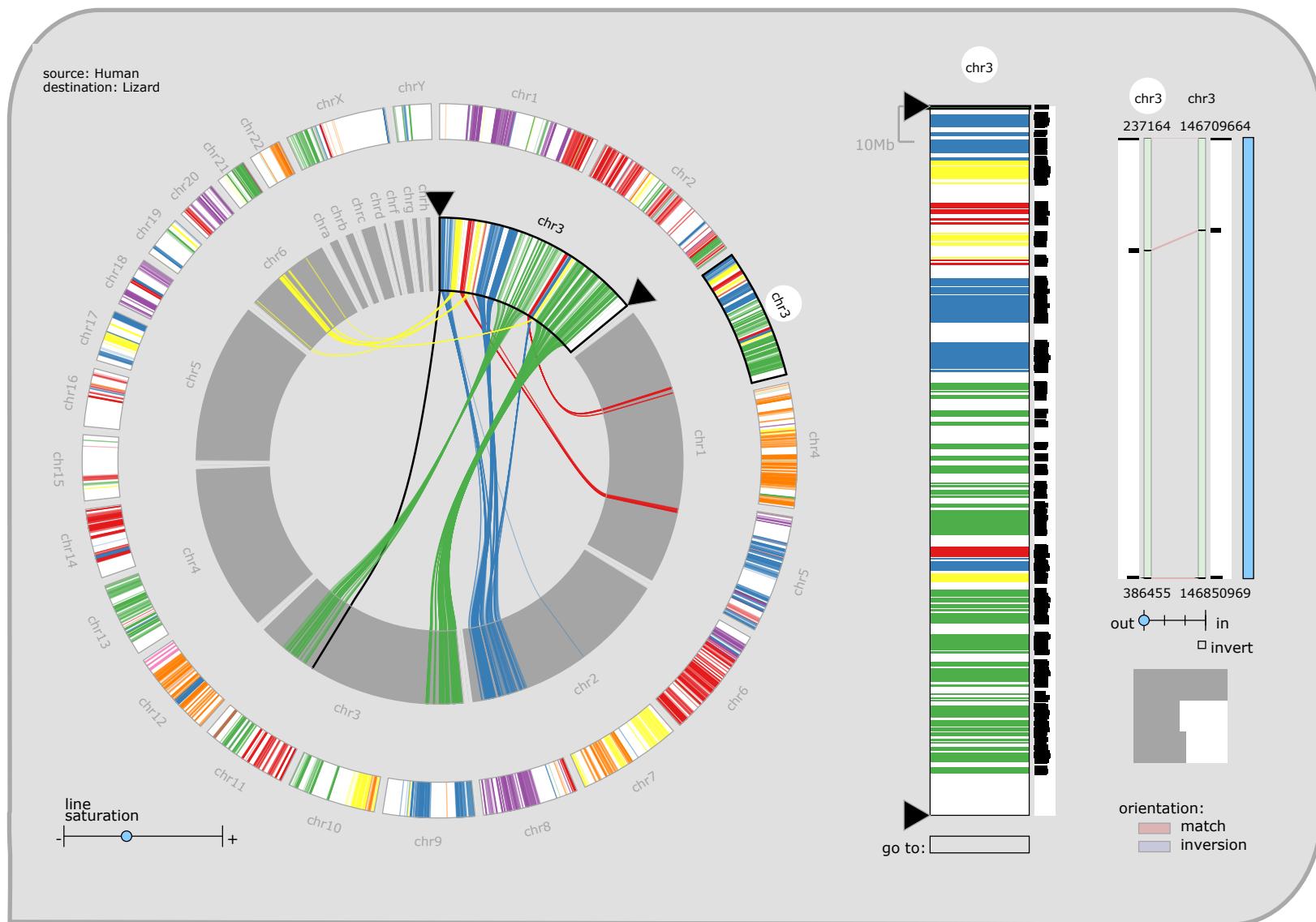


Robert Kincaid
(Agilent)



MizBee: Comparing genomes between species

- driving problems: Broad Inst. biologists studying fungus (Ma) and stickleback/pufferfish (Grabherr)
- two use phases: first fully validate computational pipeline, then can analyze biological questions
- investigated whole-genome duplication events, refined synteny block construction algorithm



MizBee <https://youtu.be/86p7brwuz2g>

Hanspeter Pfister
(Harvard)

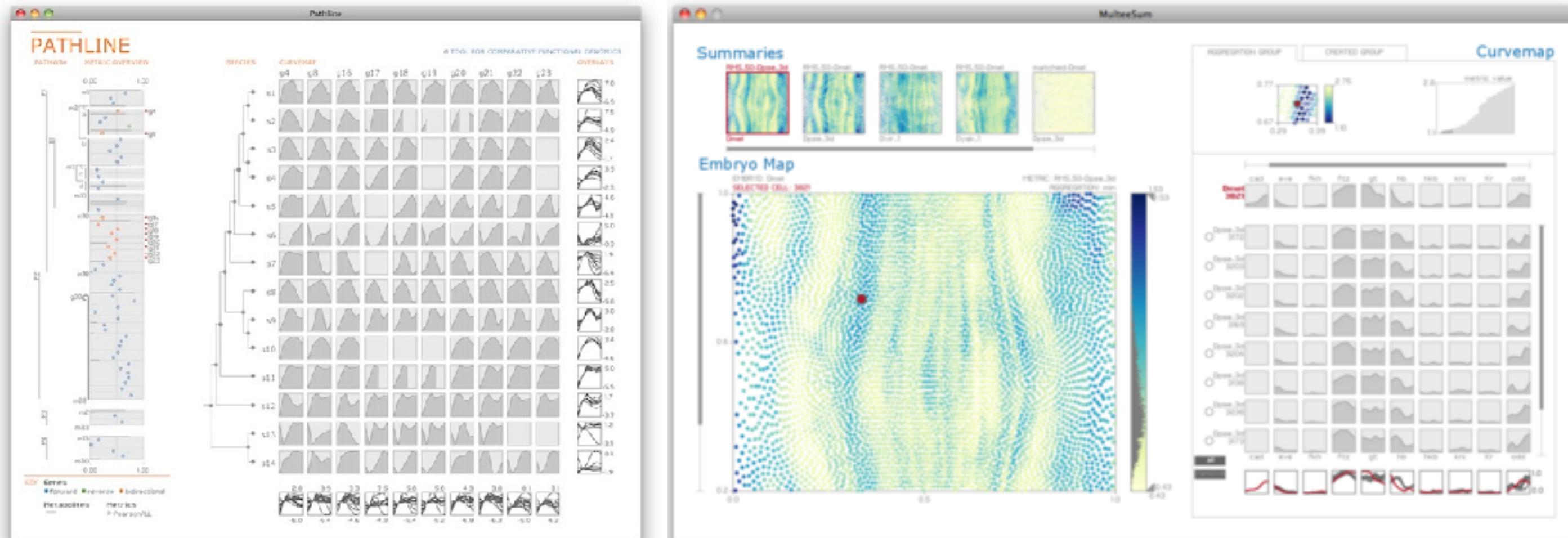


Miriah Meyer



Comparative functional genomics

- Pathline: multiple pathways, multiple genes, multiple species - over time
 - Broad Institute, Regev Lab
 - curvemap as alternative to heatmap
- MulteeSum: all that + spatial location (cells within fruitfly embryo)
 - Harvard Med School, dePace Lab
 - compare summaries across multiple computational workflows



joint work with: Meyer, Pfister, Wong, Styczynski, dePace

Variant View: Visualizing sequence variants in genetic context

- concise overview supports reasoning about variant type & location
 - across several levels of biological context (vs extensive navigation w/ genome browsers)

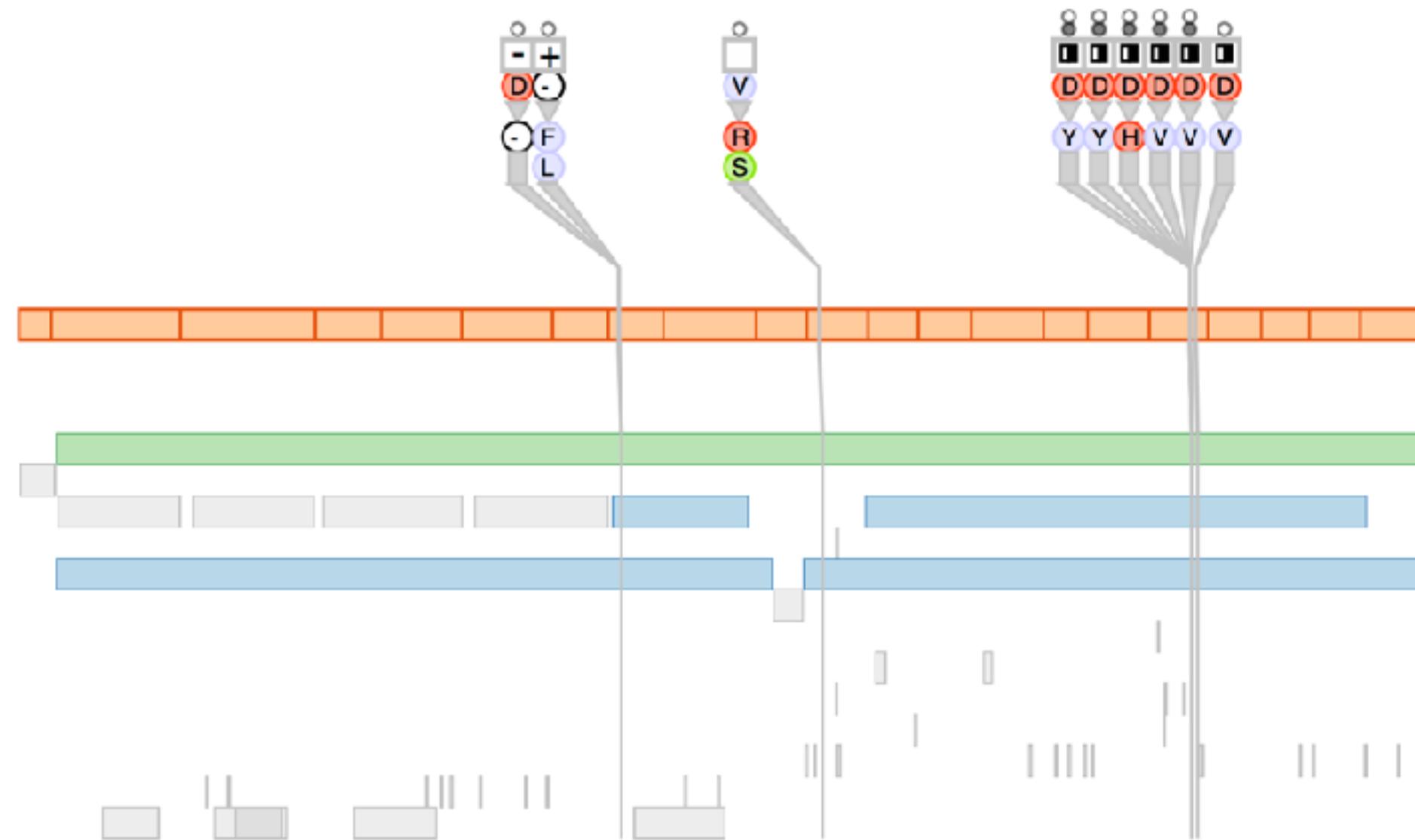
Variants
Mutation Type
Reference A.A.s
Variant A.A.s

Transcript

trans-anon

Protein

A.A. Chain Signals
Domains
Regions
Topo. Domains
Transmem.
Active Sites
NP Binding
Metal Bind.
Bindings
Mod. Residue
Carbohyd.
Disulf.



Joel Ferstay



Cydney Nielsen
(BC Cancer)



Aggregated Dendograms: Visual comparison between many phylogenetic trees

- concisely summarize trees interactively wrt bio meaningful criteria
 - one use case: compare gene trees to species trees



Zipeng Liu



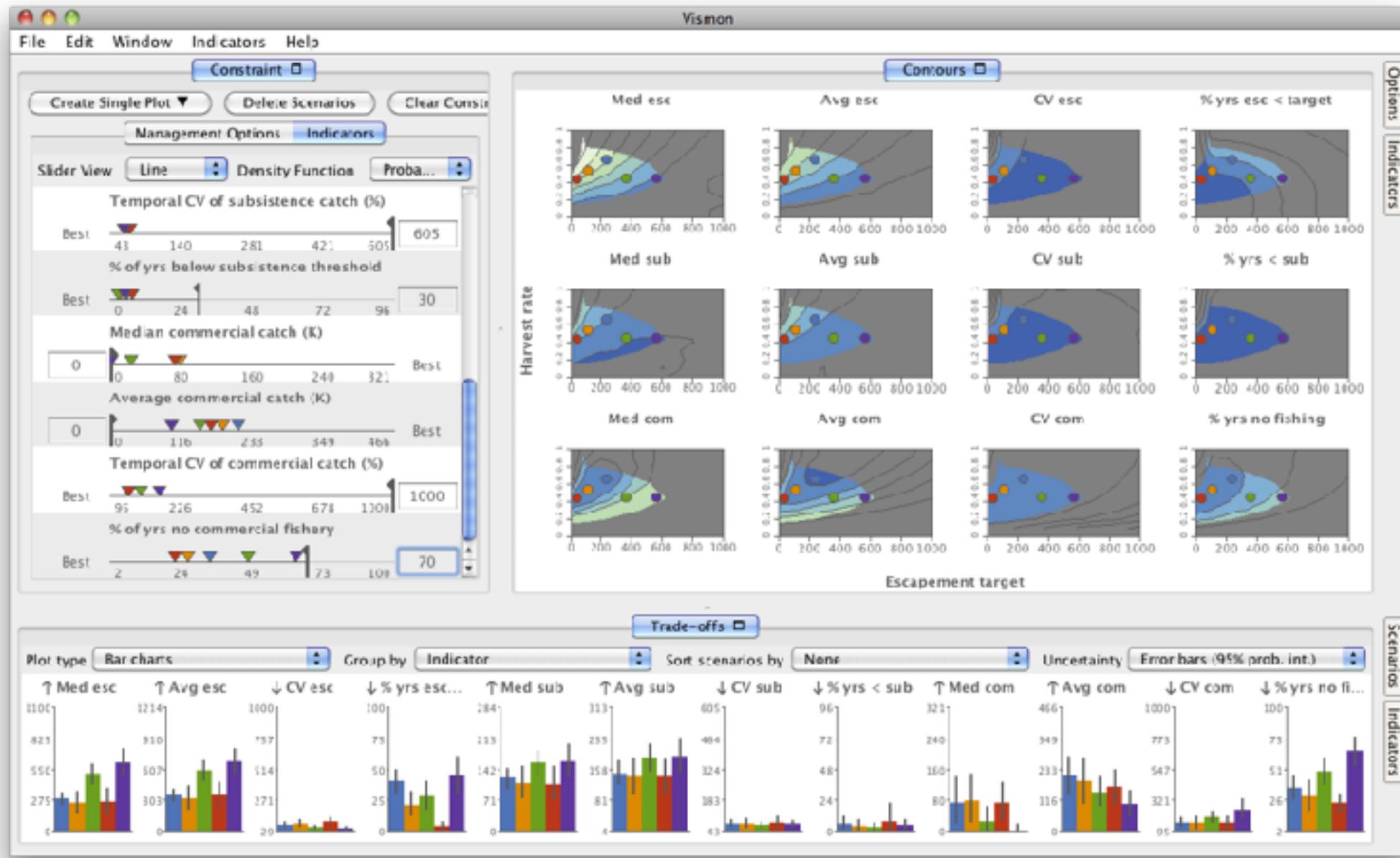
Aggregated Dendograms <https://youtu.be/2SLcz7KNLJw>

Shing Hei Zhan



Vismon: Fisheries simulation

- supporting decision-makers not expert in simulation & stats
 - sensitivity analysis, global trade-offs analysis, staged uncertainty



Vismon <https://youtu.be/h0kHoS4VYmk>

Maryam Booshehrian



Torsten Moeller (SFU)

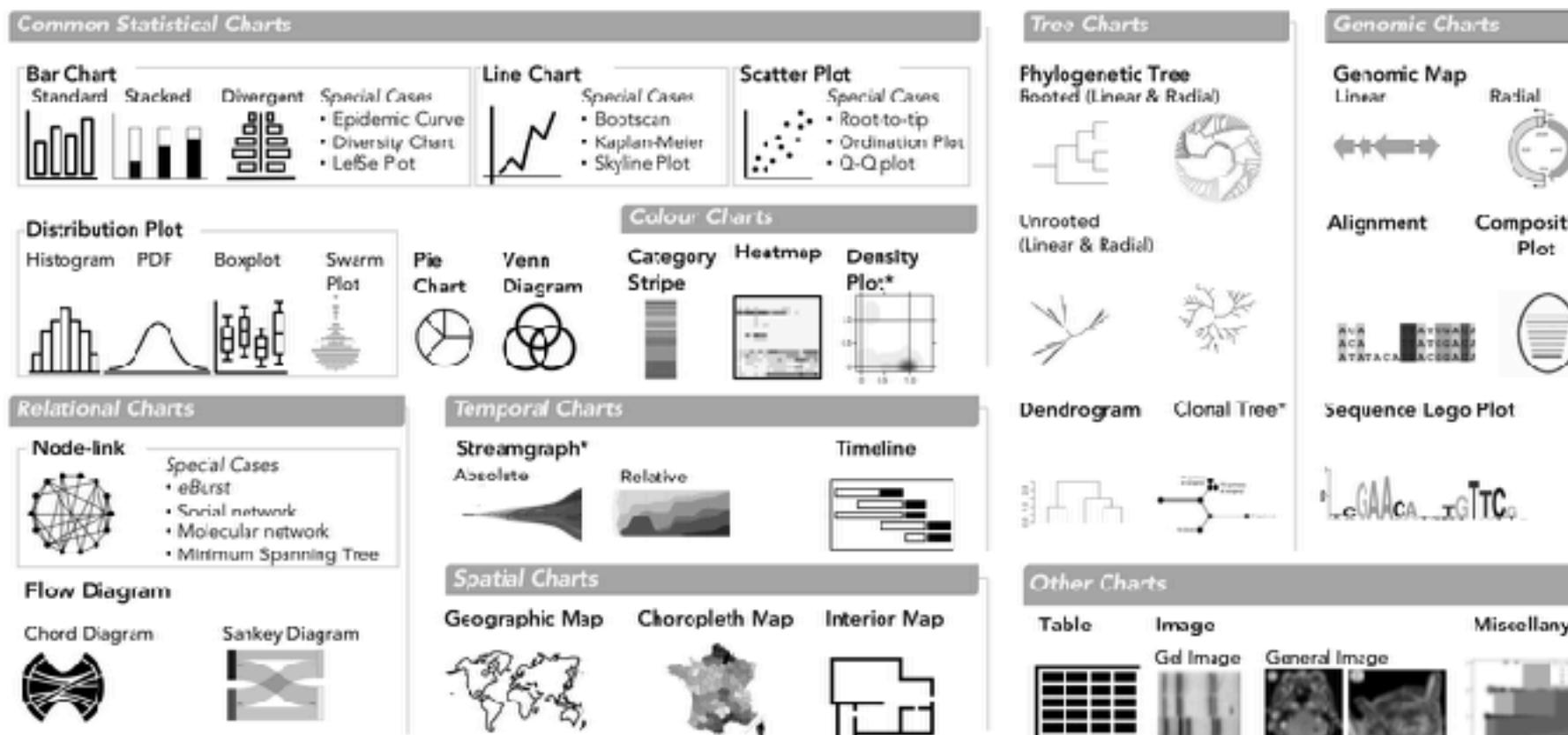


Integrating visualization & biostats methods

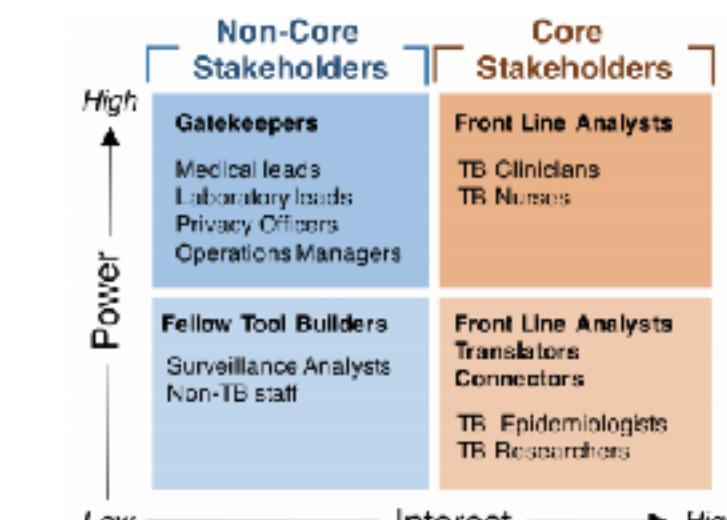
Anamaria
Crisan



- Human-centered design & qualitative coding
- Epidemiology/health expectations & constraints
- Mixed initiative: automation and manual analysis
- Mixed methods: when to use qual & when to use quant



GEViT: Genomic Epidemiology Visualization Typology
<https://gevit.net>



Regulatory & Organizational Constraints

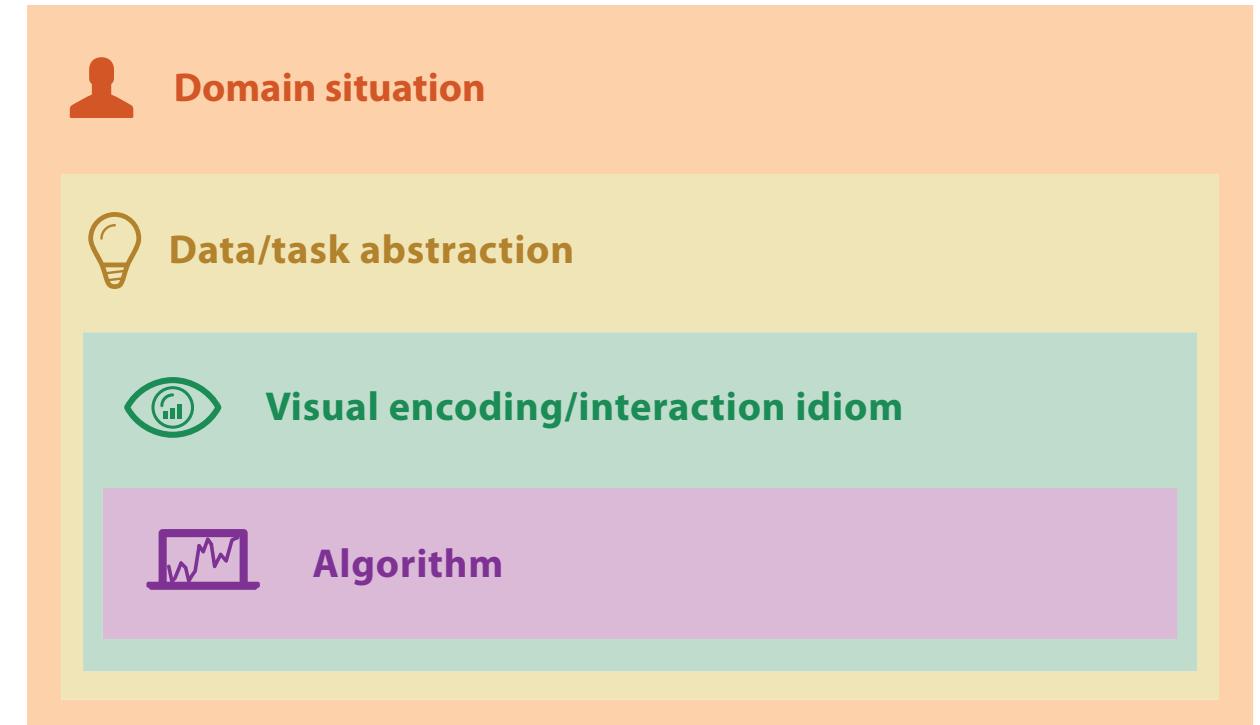
Jenn Gardy
BCCDC/SPPH



Evidence-Based Design and Evaluation of a Whole Genome Sequencing Clinical Report for the Reference Microbiology Laboratory

A Nested Model for Visualization Design and Validation

<http://www.cs.ubc.ca/labs/imager/tr/2009/NestedModel>

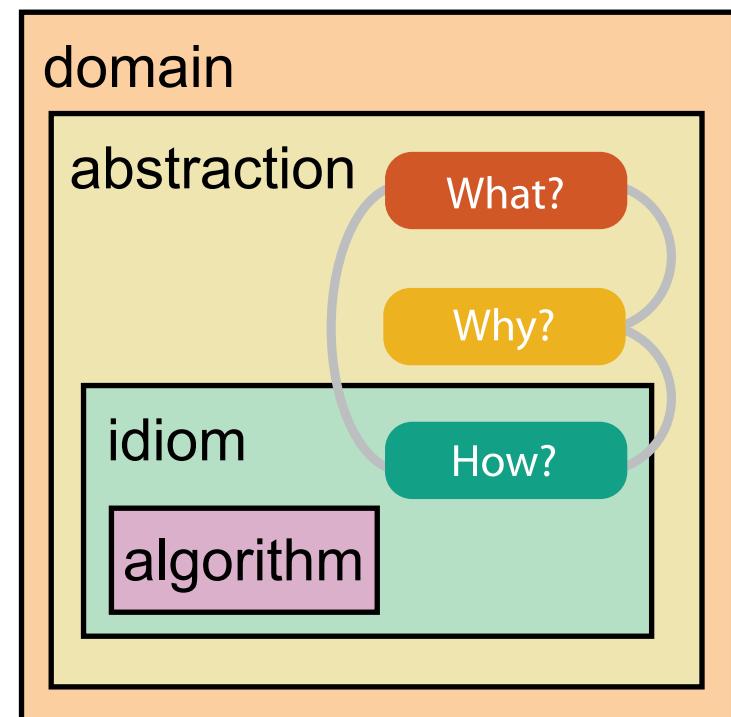


A Nested Model for Visualization Design and Validation.
Munzner. IEEE Trans. Visualization and Computer Graphics (Proc. InfoVis 09), 15(6):921-928, 2009.

Nested model: Four levels of visualization design

- *domain situation*
 - who are the target users?
- *abstraction*
 - translate from specifics of domain to **vocabulary** of visualization
 - **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).]

Different threats to validity at each level



Domain situation

You misunderstood their needs



Data/task abstraction

You're showing them the wrong thing



Visual encoding/interaction idiom

The way you show it doesn't work



Algorithm

Your code is too slow

Interdisciplinary: need methods from different fields at each level

- mix of qual and quant approaches (typically)

anthropology/
ethnography

design

computer
science

psychology

anthropology/
ethnography

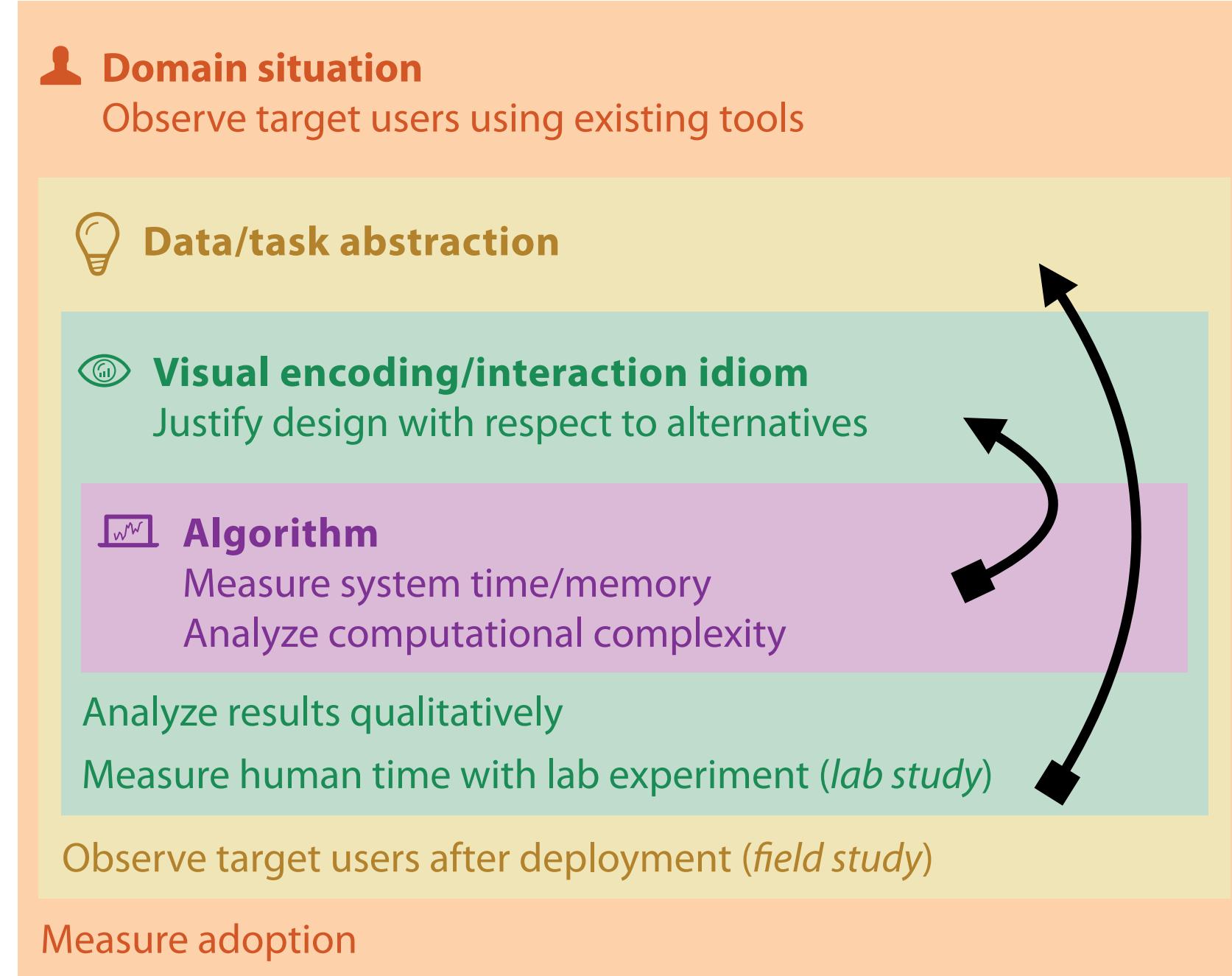
<p>👤 Domain situation Observe target users using existing tools</p>	qual
<p>💡 Data/task abstraction</p>	
<p>👁️ Visual encoding/interaction idiom Justify design with respect to alternatives</p>	qual
<p>💻 Algorithm Measure system time/memory Analyze computational complexity</p>	quant
<p>Analyze results qualitatively</p>	qual
<p>Measure human time with lab experiment (<i>lab study</i>)</p>	quant
<p>Observe target users after deployment (<i>field study</i>)</p>	qual
<p>Measure adoption</p>	quant



problem-driven
work

technique-driven
work

Mismatches: Common problem

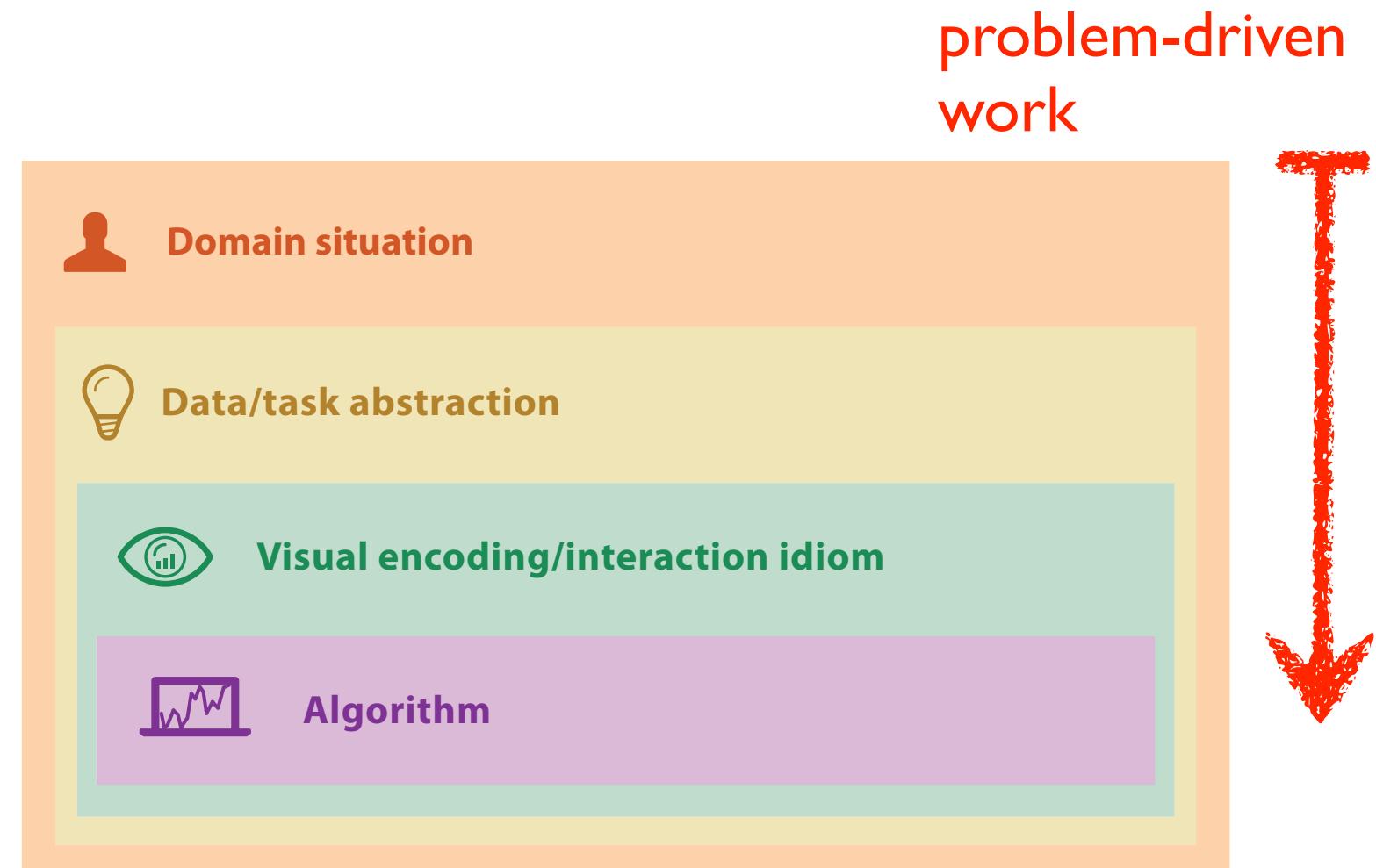


benchmarks can't
confirm design

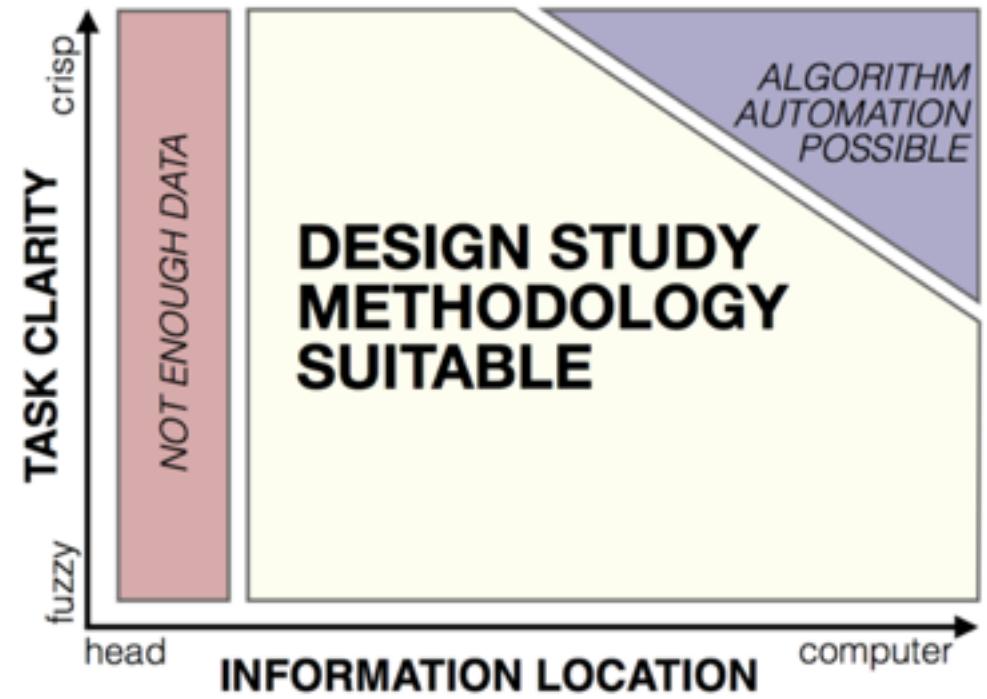
lab studies can't
confirm task
abstraction

Problem-driven collaborations

- working with domain scientists
- translating from domain-specific language
 - how to pull this off?



Building Rewarding Collaborations



Michael Sedlmair



Miriah Meyer



Design Study Methodology

Reflections from the Trenches and from the Stacks

Tamara Munzner
@tamaramunzner



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

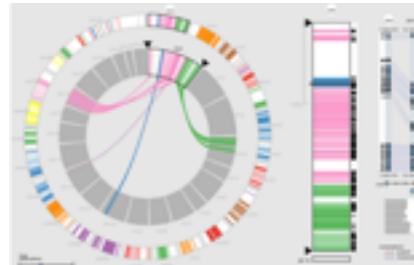
Design Study Methodology: Reflections from the Trenches and from the Stacks.

Sedlmair, Meyer, Munzner. IEEE Trans. Visualization and Computer Graphics 18(12): 2431-2440, 2012 (Proc. InfoVis 2012).

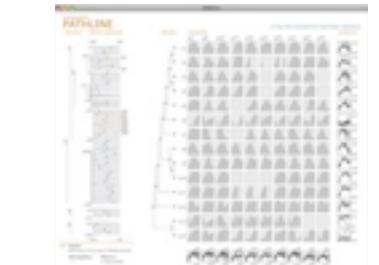
Lessons learned from the trenches: 21 between us



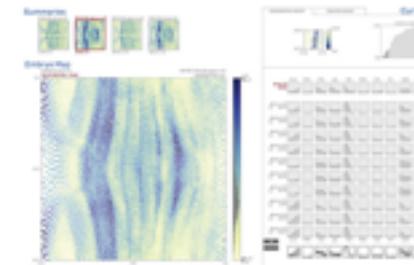
Cerebral
genomics



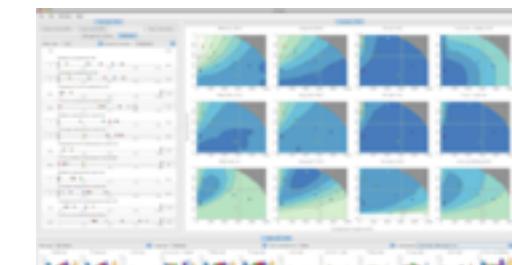
MizBee
genomics



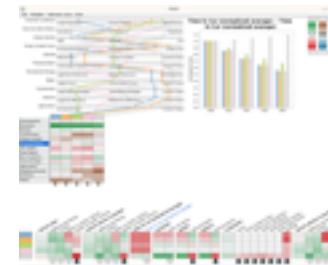
Pathline
genomics



MulteeSum
genomics



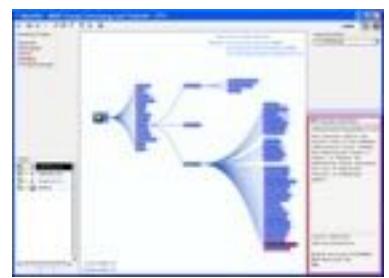
Vismon
fisheries management



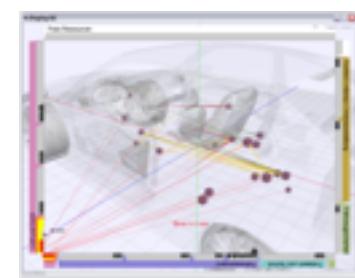
QuestVis
sustainability



WiKeVis
in-car networks



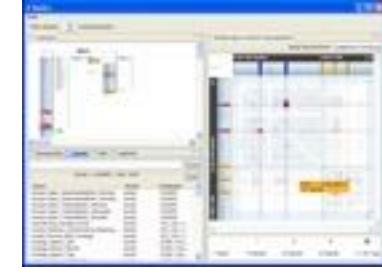
MostVis
in-car networks



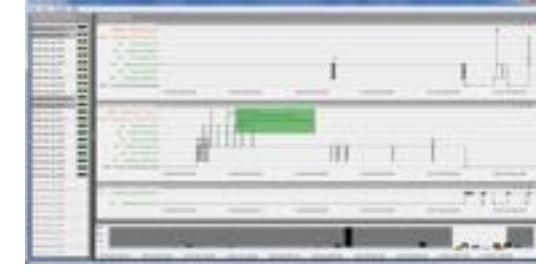
Car-X-Ray
in-car networks



ProgSpy2010
in-car networks



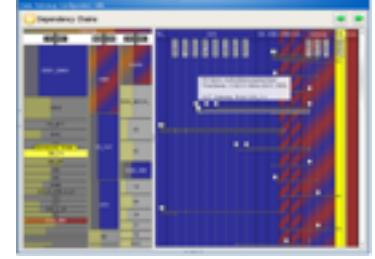
ReIEx
in-car networks



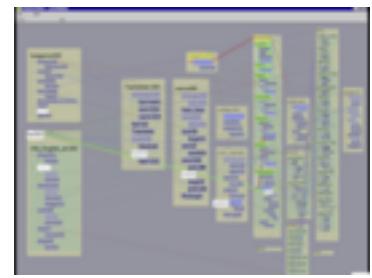
Cardiogram
in-car networks



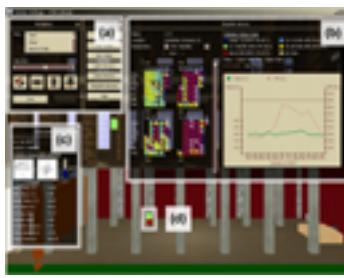
AutobahnVis
in-car networks



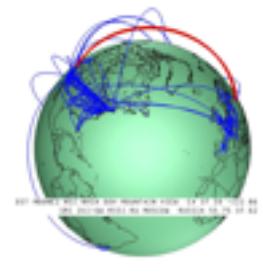
VisTra
in-car networks



Constellation
linguistics



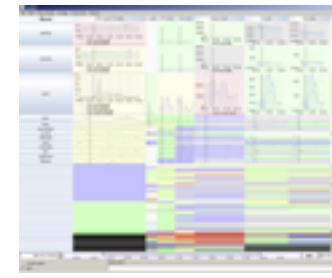
LibVis
cultural heritage



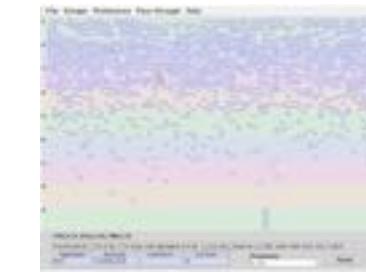
Caidants
multicast



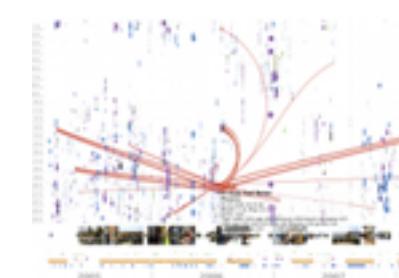
SessionViewer
web log analysis



LiveRAC
server hosting



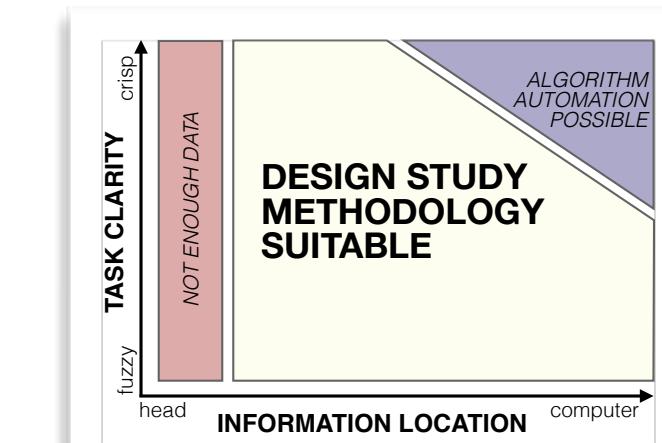
PowerSetViewer
data mining



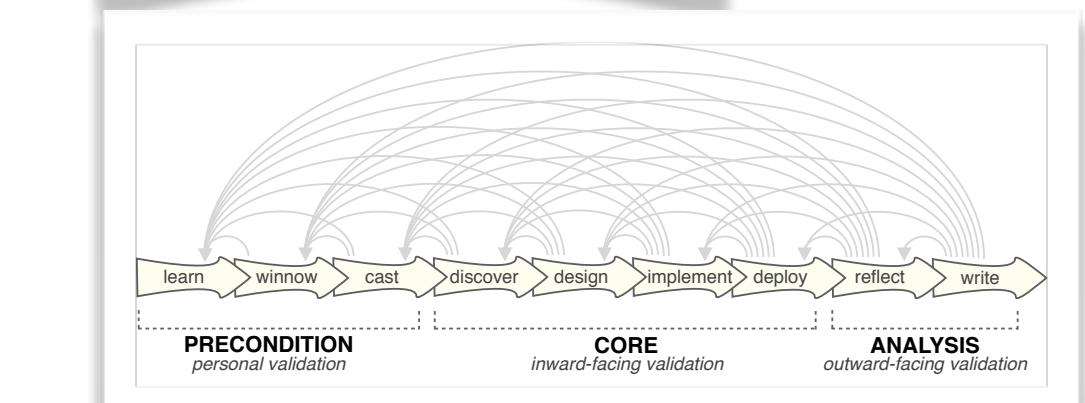
LastHistory
music listening

Methodology for problem-driven work

- definitions



- 9-stage framework



- 32 pitfalls & 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

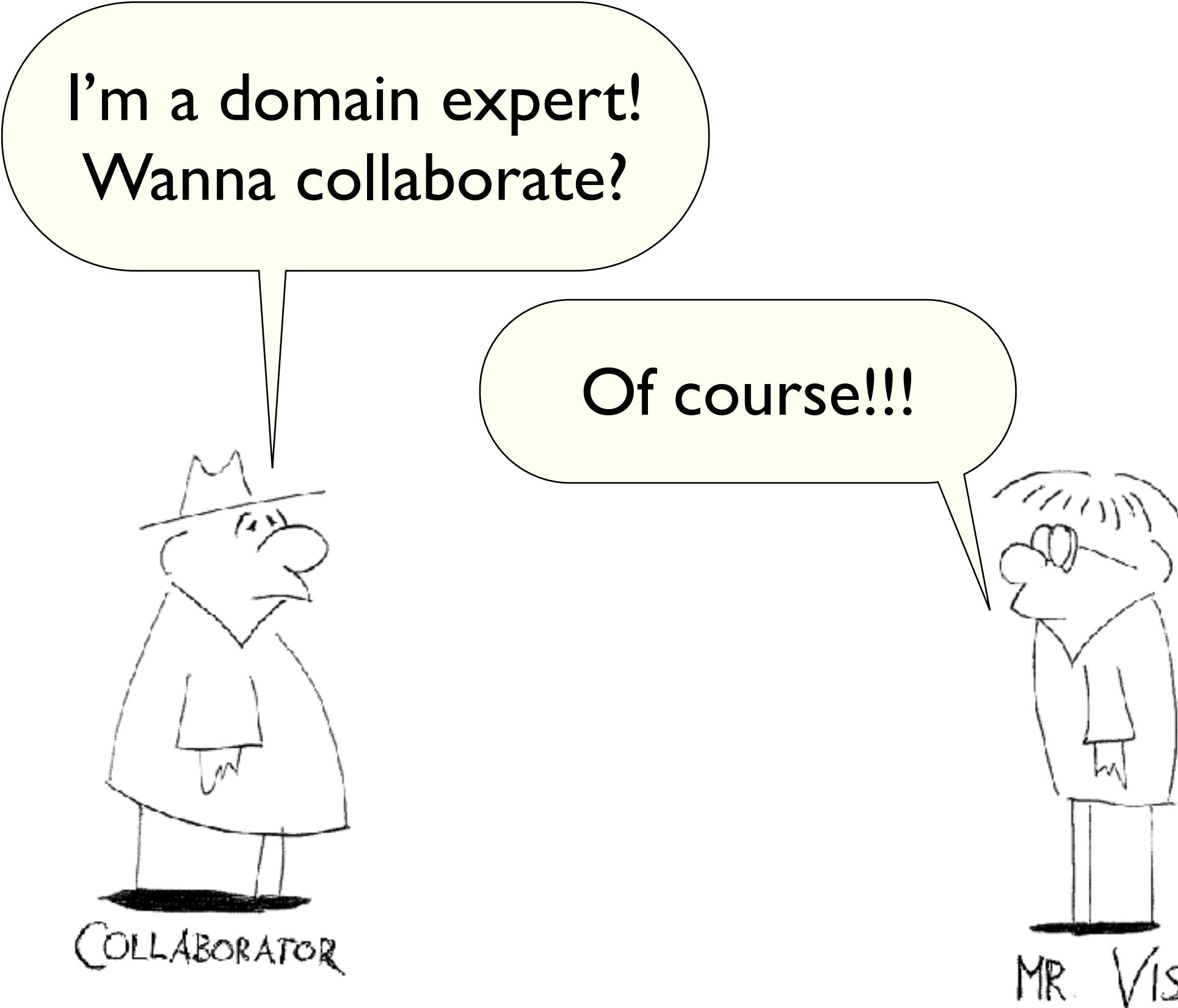
- comparison to related methodologies



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



I'm a domain expert!
Wanna collaborate?

Of course!!!

COLLABORATOR

MR. VIS

considerations



COLLABORATOR



MR. VIS



roles



COLLABORATOR

Are you a
target
user???

... or maybe a
**fellow tool
builder?**



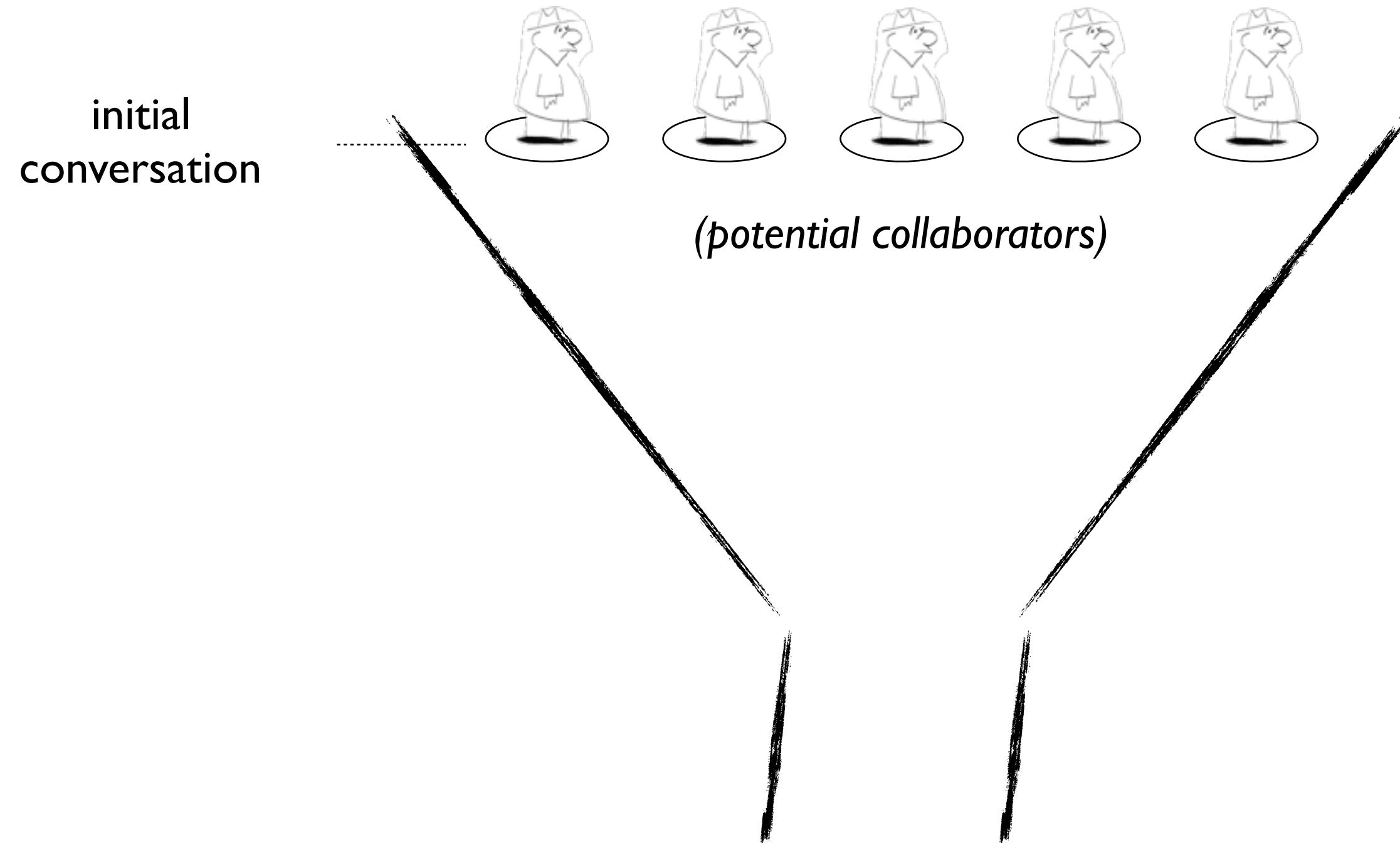
MR. VIS

METAPHOR

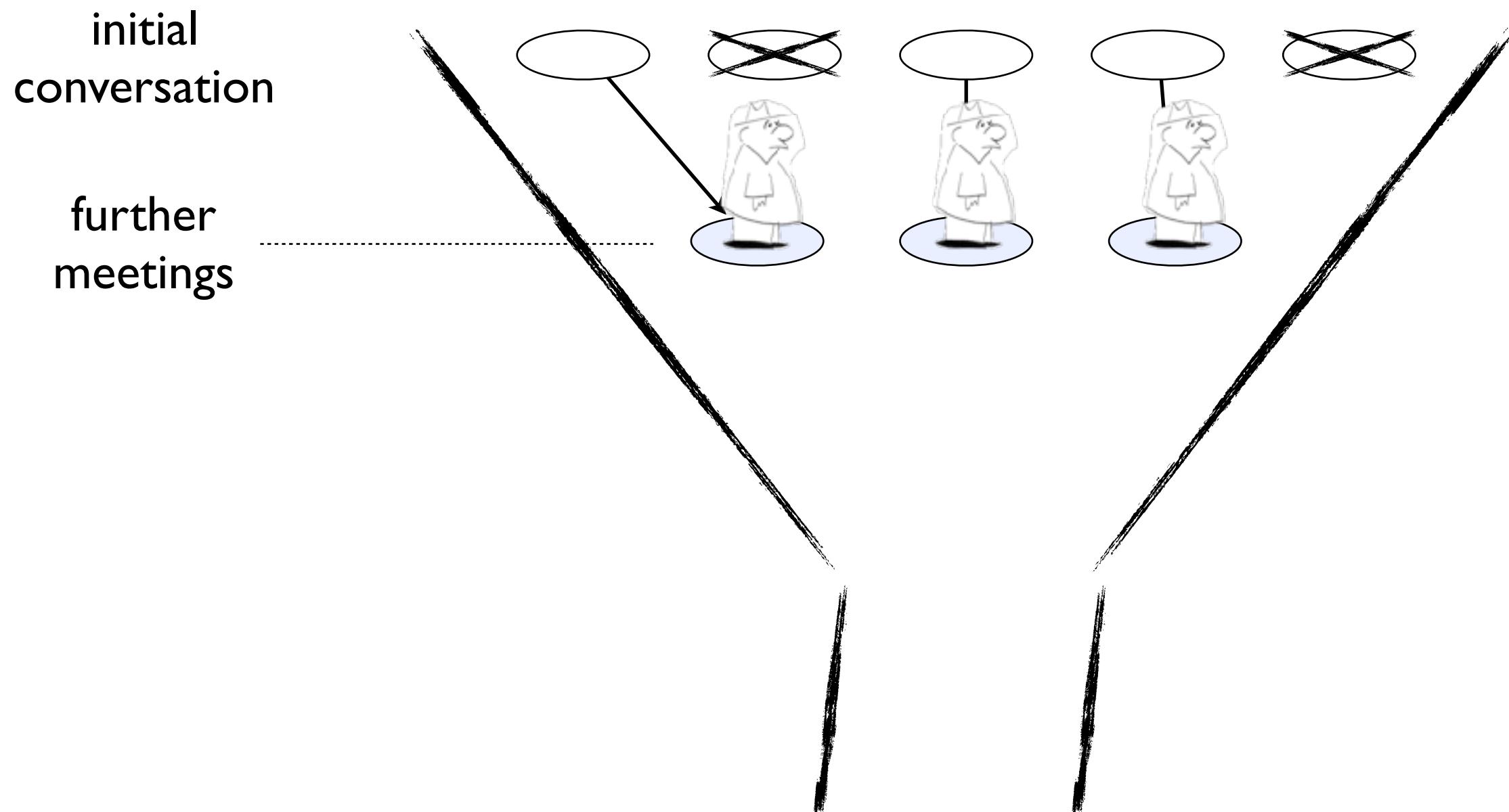
Winnowing



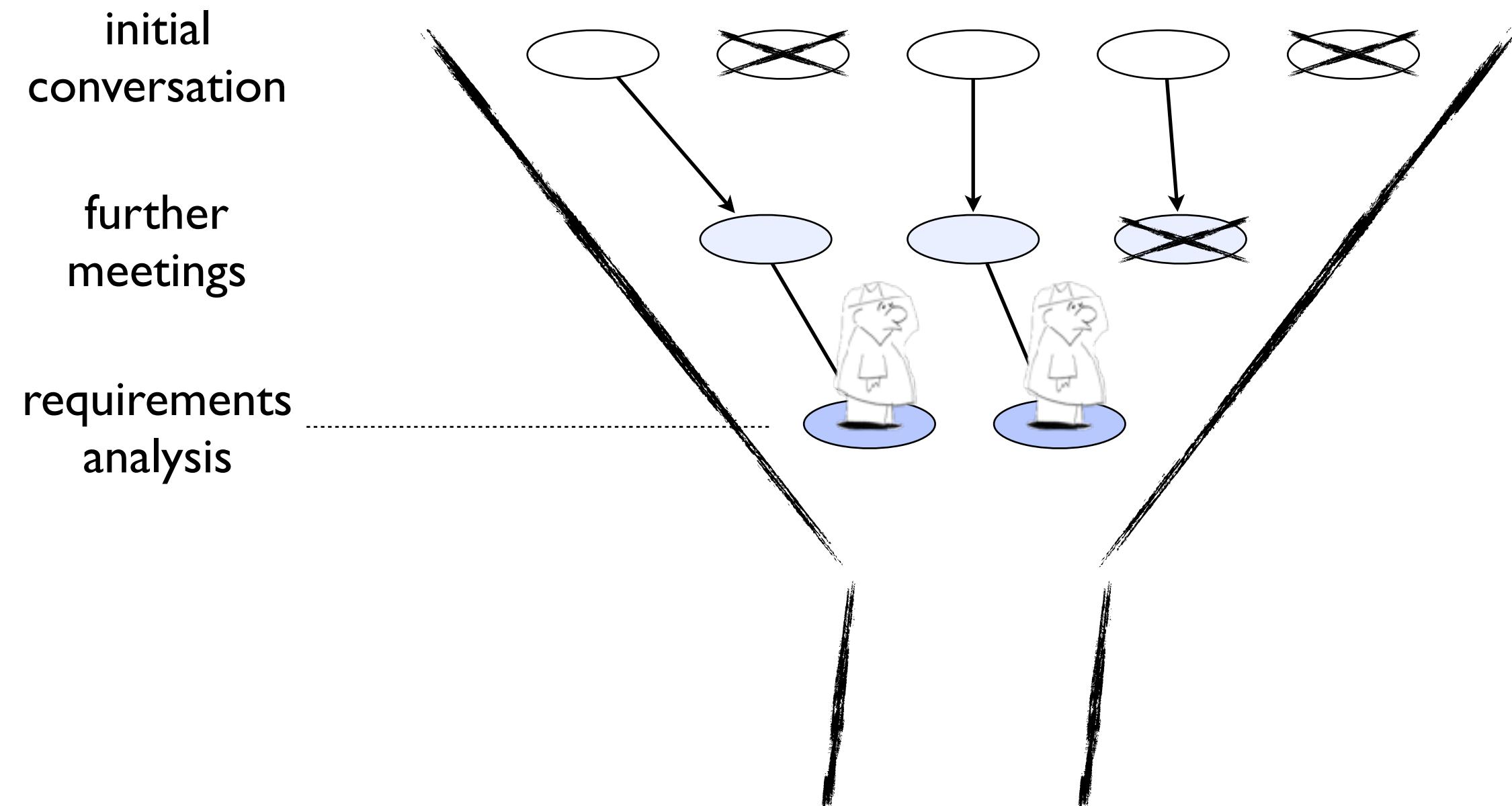
Collaborator winnowing



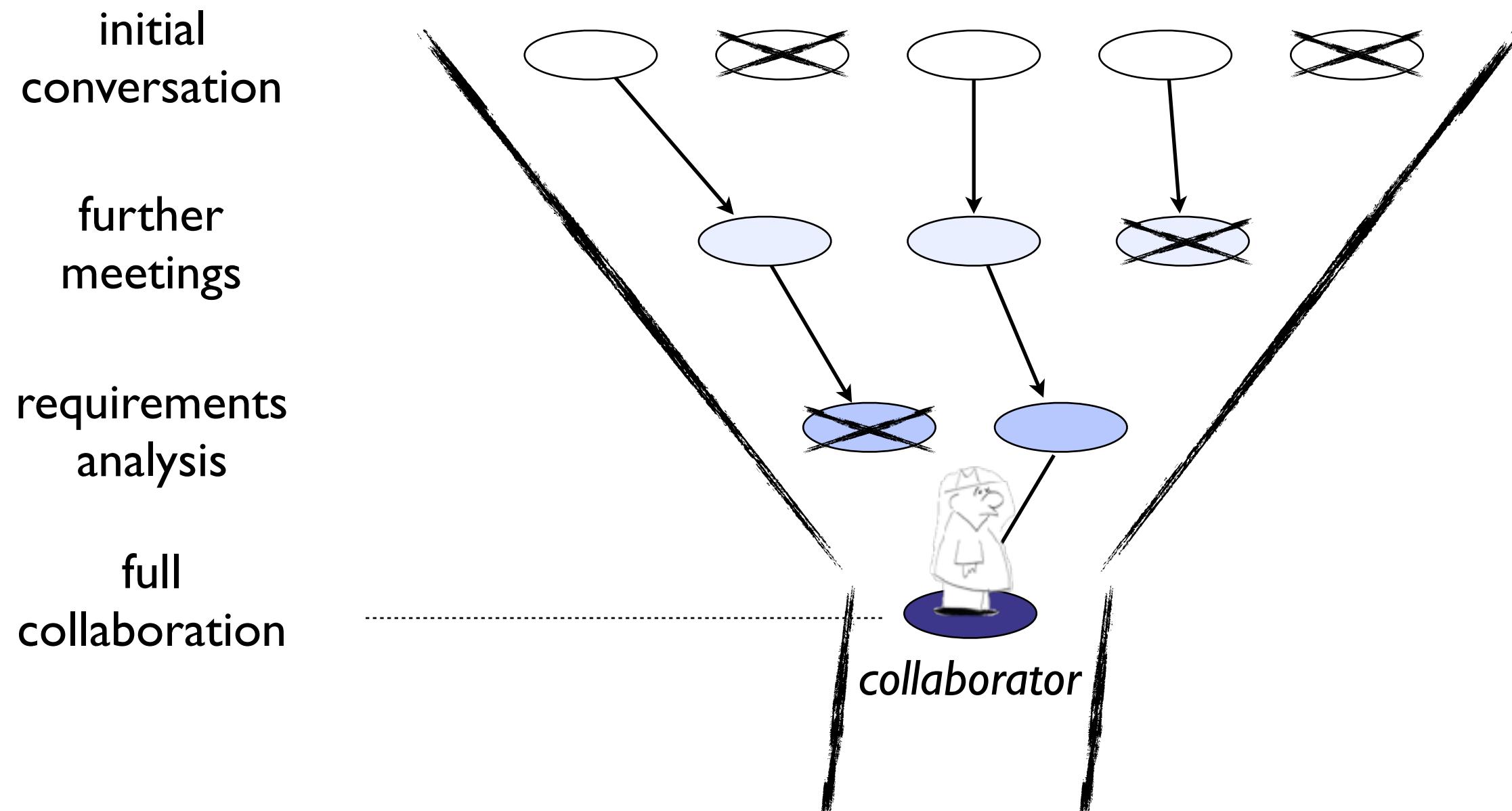
Collaborator winnowing



Collaborator winnowing



Collaborator winnowing



Collaborator winnowing

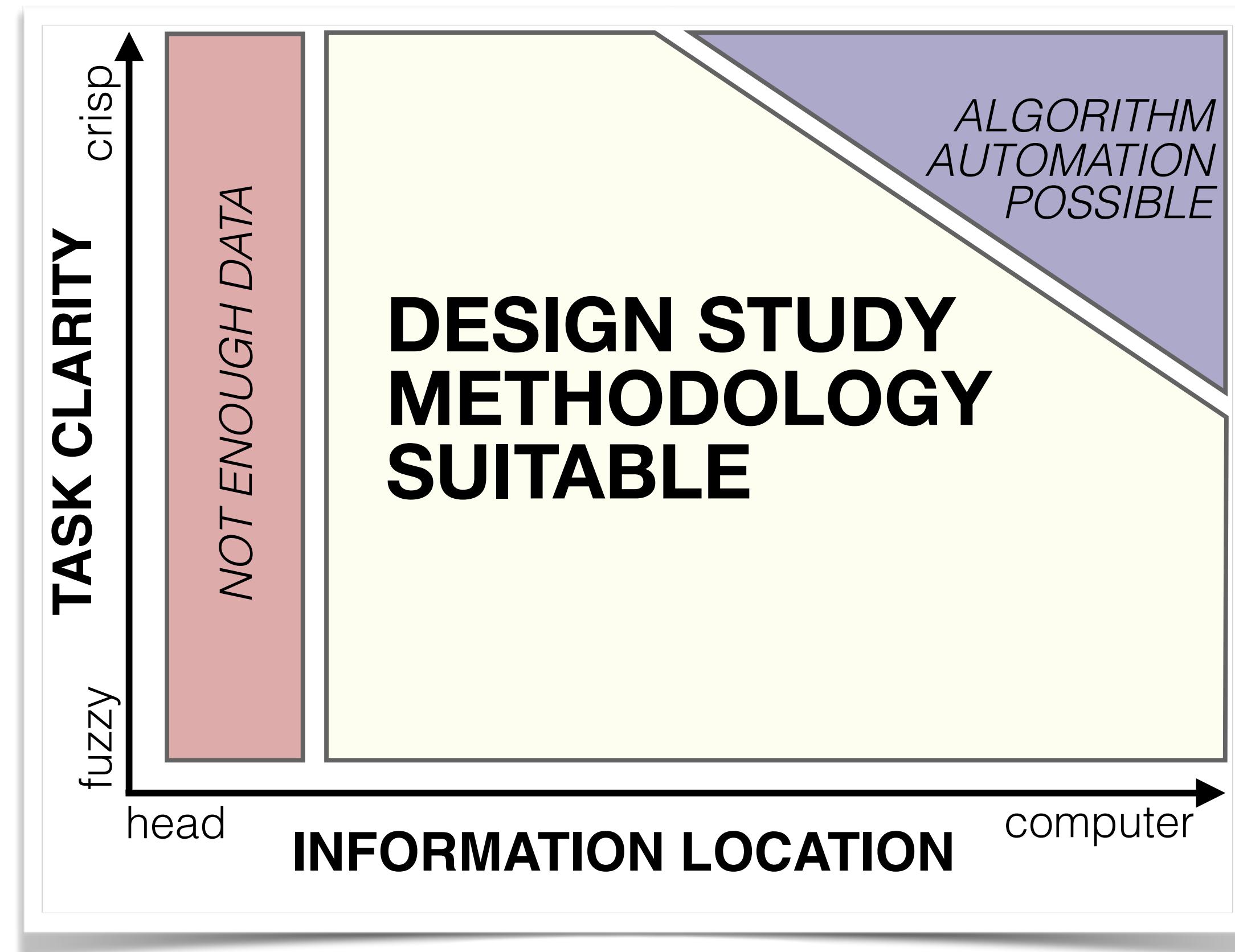


Design study methodology: 32 pitfalls

- 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

Design study methodology: definitions



More Information

- this talk
<https://www.cs.ubc.ca/~tmm/talks.html#hakai19-methods>
- papers, videos, software, talks, courses
<http://www.cs.ubc.ca/group/infovis>
<http://www.cs.ubc.ca/~tmm>

MEMBERS



Tamara Munzner



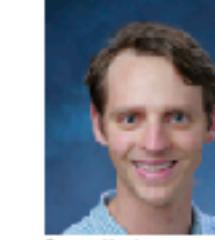
Anamaria Crisan



Zipeng Liu



Michael Oppermann



Steve Kasica



Shannah Fisher

RECENT NEWS

2/2019 [news]:

Aggregated dendograms for visual comparison between many phylogenetic trees

by Zipeng Liu, Sheng Hei Chan, and Tamara Munzner was accepted at *IEEE Transactions on Visualization and Computer Graphics*

[pre-print PDF]



10/2018 [UOC InfoVis @ IEEE VIS 2018]:

UBC InfoVis @ IEEE VIS 2018

At the *deLV Workshop*, Anamaria Crisan and Madison Elliott will present the paper "How to evaluate an evaluation study? Comparing and contrasting practices in vis with those of other disciplines".

Michael Oppermann will give a talk entitled "Uncovering Spatiotemporal Dynamics from Non-Trajectory Data" at the *Urban Data Visualization Workshop*.

Tamara Munzner will participate in a panel at the *VisGuides Workshop*. We're co-hosting the (North) West Coast Party on Thursday night!

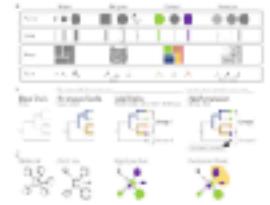


9/2018 [paper]:

A systematic method for surveying data visualizations and a resulting genomic epidemiology visualization typology: GCViT

by Anamaria Crisan, Jennifer L. Gardy, and Tamara Munzner was published in *Oxford Bioinformatics*

[paper]



8/2018 [news]:

Adjutant: an R-based tool to support topic discovery for systematic and literature reviews

by Anamaria Crisan, Tamara Munzner, and Jennifer L. Gardy was published in *Oxford Bioinformatics*

[paper]



8/2018 [highlight]:

Visiting Professor Takayuki Itoh

Takayuki Itoh from Ochanomizu University, Japan, was visiting our group between Jul - Aug 2018. Thanks for your visit, it was great having you here!



05/2018 [news]:

GoRSViz: Improving the Predicting of Self-interruption during Reading using Gaze Data

by Ian Pilzer, Shaeemah Mahmud, Vanessa Pihlman, and Tamara Munzner was accepted to *ETVIS 2018*.

[pre-print PDF]

