

Wrapup: Research Papers and Process

Tamara Munzner

Department of Computer Science
University of British Columbia

CPSC 547, Information Visualization

3 December 2020

<http://www.cs.ubc.ca/~tmm/courses/547-20>

Today

- final presentations
- final reports
 - course paper vs research paper expectations
- [evaluations]
- writing infovis papers: pitfalls to avoid
- other research pitfalls and process
 - review reading, review writing, conference talks
- next steps
 - ways to continue on with visualization

Final Presentations

Final Presentations Schedule

- 3:00-3:10 Albina Gibadullina
Geographic-Financial.
- 3:10-3:22 Alex Trostanovsky and Nikola Cucuk.
UCoD - Simplifying Supply Chain Structures in the Browser.
- 3:22-3:34 Alireza Iranpour and Jose Carvajal and Lucca Siaudzionis.
Country vs. Country: Food & Allergy Edition.
- 3:34-3:46 Anika Sayara and Namratha Rao and Roger Yu-Hsiang Lo.
Visualizing Linguistic Diversity in Vancouver.
- 3:46-3:58 Braxton Hall and Jonathan Chan and Paulette Koronkevich.
Visualizing Compiler Passes with FirstPass.
- 3:58-4:10 BREAK
- 4:10-4:22 Claude Demers-Belanger and Sanyogita Manu.
EnergyFlowVis: Visualizing Energy Use Flows for UBC Campus.
- 4:22-4:34 Cloris Feng and Derek Tam and Tae Yoon Lee.
Disease Outbreak Radar: A Tool for Epidemiologists.
- 4:34-4:44 Eric Easthope.
Bewilder: Handling Web Resource Complexity in Online Learning/Research.
- 4:44-4:56 Frank Yu and James Yoo and Lily Bryant.
Visualizing Mobility and COVID-19.
- 4:56-5:08 Gabby Xiong and Michael Cao.
Android App Similarity Visualization.
- 5:08-5:18 BREAK
- 5:18-5:30 Hannah Elbaggari and Preeti Vyas and Roopal Singh Chabra and Rubia Reis Guerra.
Firest: Visualizing the Current State and Impact of Wildfires Across Canada.
- 5:30-5:42 Huancheng Yang and Nikhil Prakash.
Smart Intersection Vis.
- 5:42-5:52 Ivan Gill.
AMR-TV: Antimicrobial Resistance Transmission Visualizer.
- 5:52-6:02 Joshua Yi Ren.
Visualizing World Color Survey Dataset
- 6:02-6:14 Kattie Sepehri and Ramya Rao Basava and Unma Desai.
Did We Save Our Tigers?
- 6:14-6:26 Raghav Goyal and Shih-Han Chou and Siddhesh Khandelwal.
README: A Literature Survey Assistant.

Final presentations

- **structure**
 - pre-created videos streamed (like pitches)
 - live Q&A
- **context**
 - CS department will be invited, also feel free to invite others
 - Piazza post with timings & zoom info
 - note different zoom URL than main class sessions
 - two short breaks
 - order: alphabetical by first name
- **code freeze**
 - no additional work on project after presentation deadline
 - additional three days to get it all written down coherently for final report

Final presentations: Thu Dec 10 3-6:30 by zoom

- length (16 projects)
 - livestreamed from my laptop: 10 min videos for groups, 8 min for solo
 - live Q&A through zoom: 2 min per project
- session structure
 - order alphabetical by first name, as on project page
 - 2 breaks, between each set of 5-6 presentations
 - dept invited, friends/others welcome
- video presentation structure
 - motivation/framing, project, results, critique/limitation
 - slides required for main part (*remember slide numbers!*)
 - demo strongly encouraged
 - should be standalone
 - don't assume audience has read proposal or updates (or remembers your pitch)
- slides/video upload
 - upload to Canvas Assignments: Final Videos, Final Slides
 - by noon Thu Dec 10

Final presentations marking

- template (may change)
 - Intro/Framing: 20%
 - Main: 30%
 - Limitations/Critique/Lessons: 10%
 - Slides: 10%
 - Presentation Style & Video: 10%
 - Demo: 10% (or N/A)
 - Question Handling: 10%
- marking by buckets
 - great 100%
 - good 89%
 - ok 78%
 - poor 67%
 - zero 0%

Marking: Course overall

- 50% Project, summative assessment at end
 - 15% Final Presentation
 - 25% Final Report
 - 60% Content
 - (penalty to 25% for missed Milestones, pass/fail)
 - pitch 5%, proposal 10%, update 10%
- 36% Async Discussion
 - 9 weeks, 4% per week
 - 75% own comments, 25% responses
 - almost all got full credit if submitted.
- 14% Sync: In-Class Participation
 - 12 sessions, 1% per session
 - 2% final presentations

Final Reports

Final reports

- PDF, use InfoVis templates http://junctionpublishing.org/vgtc/Tasks/camera_tvccg.html
 - your choice to use Latex/Word/whatever
- no length cap: illustrate freely with screenshots!
 - design study / technique: aim for at least 6-8 pages
 - analysis / survey: aim for at least 15-20 pages
- strongly encouraged to re-use text from proposal & update writeups
- encourage looking at my writing correctness and style guidelines
 - <http://www.cs.ubc.ca/~tmm/writing.html>
- strongly encourage looking at previous examples
 - www.cs.ubc.ca/~tmm/courses/547-20/projectdesc.html#examp
 - Example Past Projects (curated list)
 - direct links to all project pages to browse 2019-2003

Course requirements vs research paper standards

- research novelty **not** required
- mid-level discussion of implementation **is** required
 - part of my judgement is about how much work you did
 - high level: what toolkits etc did you use
 - medium level: what pre-existing features did you use/adapt
 - low level **not** required: manual of how to use, data structure details
- design justification **is** required
 - (unless analysis/survey project)
 - different in flavour between design study projects and technique projects
 - technique explanation alone is not enough
- publication-level validation **not** required
 - user studies, extensive computational benchmarks, utility to target audience

Report structure: General

- low level: necessary but not sufficient
 - correct grammar/spelling
 - sentence flow
- medium level: order of explanations
 - build up ideas
- high through low level: why/what before how
 - paper level
 - motivation: why should I care
 - overview: what did you do
 - details: how did you do it
 - section level
 - overview then details
 - sometimes subsection or paragraph level

Sample outlines: Design study

- www.cs.ubc.ca/~tmm/courses/547-20/projectdesc.html#examp
- abstract
 - concise summary of your project
 - do not include citations
- introduction
 - give big picture, establish scope, some background material might be appropriate
- related work
 - include both work aimed at similar problems and similar solutions
 - no requirement for research novelty, but still frame how your work relates to it
 - cover both academic and relevant non-academic work
 - you might reorder to have this section later

Sample outlines: Design study II

- data and task abstractions
 - analyze your domain problem according to book framework (what/why)
 - include both domain-language descriptions and abstract versions
 - could split into data vs task, then domain vs abstract - or vice versa!
 - typically data first then task, so that can refer to data abstr within task abstr
- solution
 - describe your solution idiom (visual encoding and interaction)
 - analyze it according to book framework (how)
 - justify your design choices with respect to alternatives
 - if significant algorithm work, discuss algorithm and data structures

Sample outlines: Design study III

- **implementation**
 - medium-level implementation description
 - specifics of what you wrote vs what existing libraries/toolkits/components do
 - breakdown of who did what work & updated milestones (actual vs estimates)
- **results**
 - include scenarios of use illustrated with multiple screenshots of your software
 - walk reader through how your interface succeeds (or falls short) of solving intended problem
 - report on evaluation you did (eg deployment to target users, computational benchmarks)
 - screenshots should be png (lossless compression) not jpg (lossy compression)!
- **discussion and future work**
 - reflect on your approach: strengths, weaknesses, limitations
 - lessons learned: what do you know now that you didn't when you started?
 - future work: what would you do if you had more time?

Sample outlines: Design study IV

- conclusions
 - summarize what you've done
 - different than abstract since reader has seen all the details
- bibliography
 - make sure to use real references for work that's been published academically
 - not just URL
 - check arxiv papers, many have forward link to final publication venue - use that too!
 - be consistent! most online sources require cleanup including IEEE/ACM DLs
 - do pay attention to my instructions for checking reference consistency
 - <http://www.cs.ubc.ca/~tmm/writing.html#refs>

Sample outlines: Technique (diffs)

- *Abstract, Introduction (same as above)*
- Related Work
 - big focus on similar solutions, some discussion of similar problems (same task/data combo)
- Data and Task Abstractions
 - much shorter than the corresponding one for design studies, framing context not core contrib
- Solution
 - describing proposed idiom exactly, not justifying its use for particular domain problem
 - as above, analyze in terms of design choices, justify why appropriate vs alternatives
- *Implementation (same as above)*
- Results
 - less emphasis on scenarios with particular target users
 - more emphasis on characterizing the breadth of possible uses
 - still definitely include screenshots of the system in action
- *Discussion / Future Work, Conclusions, Bibliography (same as above)*

Sample outlines: Survey (diffs)

- *Abstract (same as above)*
- **Introduction**
 - discuss the scope of what you're covering, why it's interesting/reasonable partition compared to visualization as a whole
- **Related Work**
 - **only** previous surveys
 - focus on how your work is similar to or different from them, especially wrt coverage
- **Main**
 - break up into sections based on your own synthesis of themes of work covered
 - you might want a Background section at the start if domain-focused survey
 - where there's important vocabulary/ideas to establish before diving into main discussion
 - analyze visualizations proposed in these papers in terms of what/why/how framework
 - include images from papers
- *Discussion / Future Work, Conclusions, Bibliography (same as above)*

Sample outlines: Analysis (diffs)

- *Abstract, Intro (same as above)*
- **Domain Background**
 - relevant vocabulary/ideas, your own background/connection
- **Data/Task Abstraction, Related Work (same as above)**
- **Methods and Tools**
 - how has it previously/normally been analyzed
 - explain what idioms you chose and justify those choices; same for tools
- **Analysis**
 - present results of your visual data analysis, including screenshots of tools in action
 - specifics of what you learned in terms of the domain problem
 - your reflection on how visualization choices helped you understand it
 - strengths/weaknesses of your approach (idioms and tools)
 - can be interleaved or in separate section at end
- *Discussion / Future Work, Conclusions, Bibliography (same as above)*

Sample outlines: Other types

- see page for implementation project types
 - implementation
www.cs.ubc.ca/~tmm/courses/547-20/projectdesc.html#outlines
- interactive explanations
 - structurally identical to design study, although actual content will differ a bit

Report marking

- required: at least material I've listed
 - you may include more material, you may choose alternate orderings
- probable marking scheme (may change!)
 - design study & technique & explainer: 12.5% each for
 - intro, related work, abstractions, solution, implementation, results, discussion, style
 - style: 10% main, 2.5% bibliography
 - survey: intro (10%), relwork (10%), main (60%), style (20%)
 - analysis: intro/domain (8%), abstr (8%), relwork (8%), methods/tools (8%), analysis (52%), discussion (8%), style (8%)
- reminder: project content is 60% of entire project mark
 - report is 25%, presentation is 15%

Code / Video

- required: submit your code
 - so I can see what you've done, but I will not post
 - include README.txt file at root with brief roadmap/overview of organization
 - which parts are your code vs libraries
 - how to compile and run
 - I do not necessarily expect your code compiles on my machine
- encouraged but not required
 - submit live demo URL (provide in README.txt file)
 - open-source your code (if so, fine to just send me that URL)
 - submit supporting video (if different from final presentation)
 - with or without voiceover
 - very nice to have later, software bitrot makes demos not last forever!

Showcase image

- showcase image for projects page
 - 300x300 image
 - call it showcase.png or showcase.jpg

Logistics

- **Assignments: Final Presentations on Canvas**
 - videos & slides upload due Thu Dec 10, noon
 - (3.5 hrs before presentation session)
- **Assignments: Final Report on Canvas**
 - upload due Mon Dec 14 8pm (PST)
 - required & posted: report, showcase image
 - required but not posted: code including README
 - encouraged: live demo URL, video (if different from final present video)

Evaluations

Process & Pitfalls for InfoVis Papers

Pitfalls

- writing infovis papers: pitfalls to avoid
 - Process and Pitfalls in Writing Information Visualization Research Papers.
Tamara Munzner. In: Information Visualization: Human-Centered Issues and Perspectives.
Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, Chris North, eds.
Springer LNCS Volume 4950, p 134-153, 2008.

Idiom pitfalls

- **Unjustified Visual Encoding**
 - should justify why visual encoding design choices appropriate for problem
 - prerequisite: clear statement of problem and encoding!
- **Hammer In Search of Nail**
 - should characterize capabilities of new technique if proposed in paper
- **Color Cacophony**
 - avoid blatant disregard for basic color perception issues
 - huge areas of highly saturated color
 - categorical color coding for 15+ category levels
 - red/green without luminance differences
 - encoding 3 separate attributes with RGB
- **Rainbows Just Like In The Sky**
 - avoid hue for ordered attribs, perceptual nonlinearity along rainbow gradient

Later pitfalls: Strategy

- **What I Did Over My Summer Vacation**
 - don't focus on effort rather than contribution
 - don't be too low level, it's not a manual
- **Least Publishable Unit**
 - avoid tiny increment beyond (your own) previous work
 - bonus points: new name for old technique
- **Dense As Plutonium**
 - don't cram in so much content that can't explain why/what/how
 - fails reproducibility test
- **Bad Slice and Dice**
 - two papers split up wrong
 - neither is standalone, yet both repeat

Later pitfalls: Tactics

- **Stealth Contributions**

- don't leave them implicit, it's your job to tell reader explicitly!
- consider carefully, often different from original project goals

Contributions in research papers

- what are your research contributions?
 - what can we do that wasn't possible before?
 - how can we do something better than before?
 - what do we know that was unknown or unclear before?
- determines everything
 - from high-level message to which details worth including
- often not obvious
 - diverged from original goals, in retrospect
- state them explicitly and clearly in the introduction
 - don't hope reviewer or reader will fill them in for you
 - don't leave unsaid should be obvious after close reading of previous work
 - goal is clarity, not overselling (limitations typically later, in discussion section)

Later pitfalls: Tactics

- **Stealth Contributions**
 - don't leave them implicit, it's your job to tell reader explicitly!
 - consider carefully, often different from original project goals
- **I Am So Unique**
 - don't ignore previous work
 - both on similar problems and with similar solutions
- **Enumeration Without Justification**
 - “X did Y” not enough
 - must say why previous work doesn't solve your problem
 - what limitations of their does your approach fix?
- **I Am Utterly Perfect**
 - no you're not; discussion of limitations makes paper stronger!

Later pitfalls: Results

- **Unfettered By Time**
 - choose level of detail for performance numbers
 - detailed graphs for technique papers, high-level for design & eval papers
- **Straw Man Comparison**
 - compare appropriately against state-of-the-art algorithms
 - head-to-head hardware is best (re-run benchmarks yourself, all on same machine)
- **Tiny Toy Datasets**
 - compare against state-of-the-art dataset sizes for technique (small ok for eval)
- **But My Friends Liked It**
 - asking labmates not convincing if target audience is domain experts
- **Unjustified Tasks**
 - use ecologically valid user study tasks: convincing abstraction of real-world use

Final pitfalls: Style

- **Deadly Detail Dump**
 - explain *how* only **after** *what* and *why*; provide high-level framing before low-level detail
- **Story-Free Captions**
 - optimize for flip-through-pictures skimming
- **My Picture Speaks For Itself**
 - explicitly walk them through images with discussion
- **Grammar Is Optional**
 - good low-level flow is necessary (but not sufficient), native speaker check good if ESL
- **Mistakes Were Made**
 - don't use passive voice, leaves ambiguity about actor
 - your research contribution or done by others?

Final pitfalls: Style 2

- Jargon Attack
 - avoid where you can, define on first use
 - all acronyms should be defined
- Nonspecific Use Of Large
 - quantify! hundreds? 10K? 100K? millions? billions?...

Final pitfalls: Submission

- **Slimy Simultaneous Submission**
 - often detected when same reviewer for both
 - instant dual rejection, often multi-conference blacklist
- **Resubmit Unchanged**
 - respond to previous reviews: often get reviewer overlap, irritated if ignored

Generality

- encoding: visualization specific
- strategy: all research
- tactics: all research
- results: visualization specific
- style: all research, except
 - Story-Free Captions, My Picture Speaks For Itself

Research Process & Pitfalls

Review reading pitfalls

- Reviewers Were Idiots
 - rare: insufficient background to judge worth
 - if reviewer didn't get your point, many readers won't
 - your job: rewrite so clearly that nobody can misunderstand
- Reviewers Were Threatened By My Brilliance
 - seldom: unduly harsh since intimately familiar with area
- I Just Know Person X Wrote This Review
 - sometimes true, sometimes false
 - don't get fixated, try not to take it personally
- It's The Writing Not The Work
 - sometimes true: bad writing can doom good work (good writing may save borderline)
 - sometimes false: weak work common! reinvent the wheel worse than previous one

Review writing pitfalls

- **Uncalibrated Dismay**
 - remember you've only read the best of the best!
 - most new reviewers are overly harsh
- **It's Been Done, Full Stop**
 - you must say who did it in which paper, full citation is best
- **You Didn't Cite Me**
 - stop and think whether it's appropriate
 - be calm, not petulant
- **You Didn't Channel Me**
 - don't compare against paper you would have written
 - review the paper they submitted

Conference talk pitfalls

- **Results As Dessert**
 - don't save until the end as a reward for the stalwart!
 - showcase early to motivate
- **A Thousand Words, No Pictures**
 - aggressively replace words with illustrations
 - most slides should have a picture
- **Full Coverage Or Bust**
 - cannot fit all details from paper
 - communicate big picture
 - talk as advertising: convince them it's worth their time to read paper!

Paper writing process suggestions

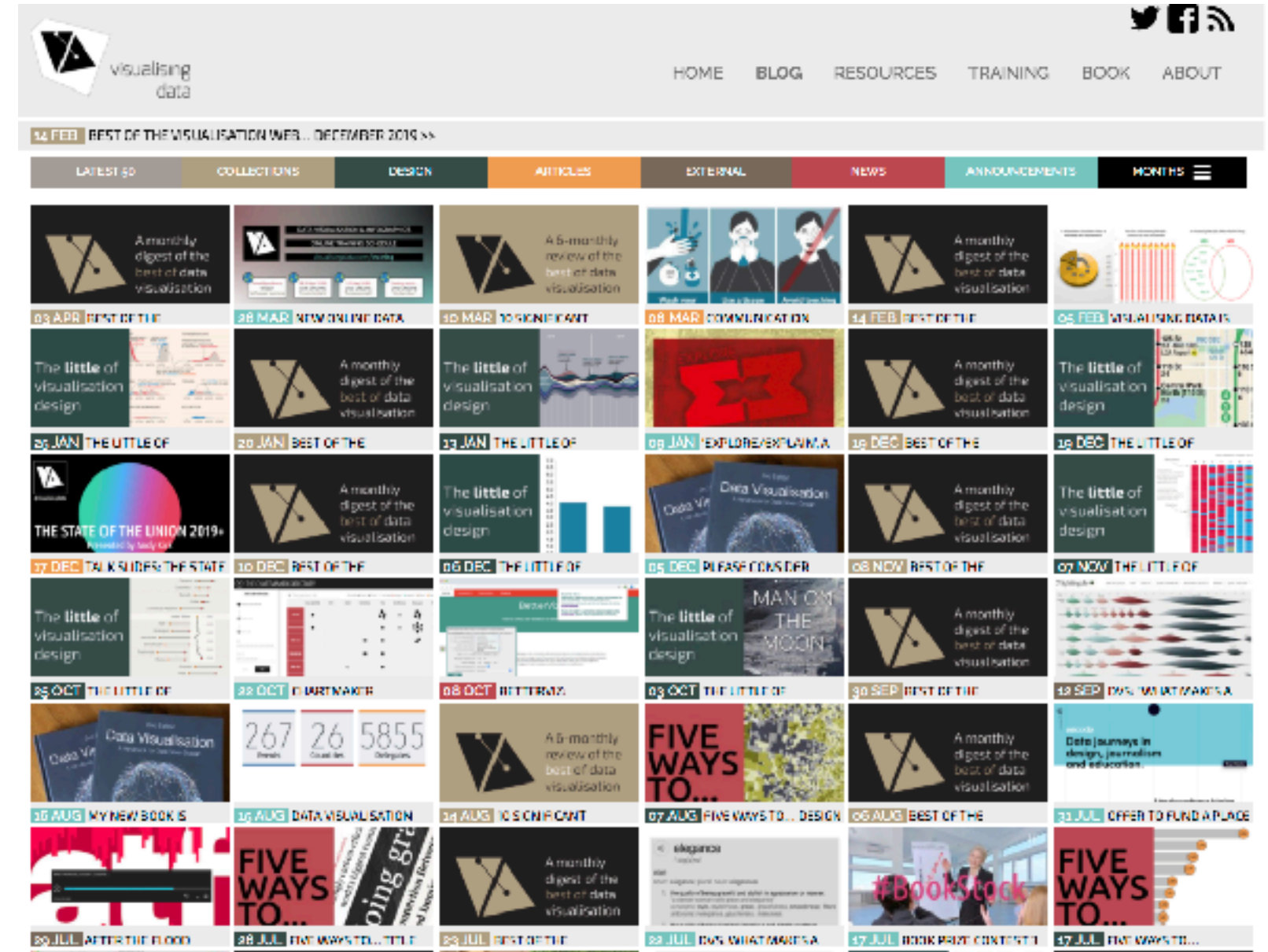
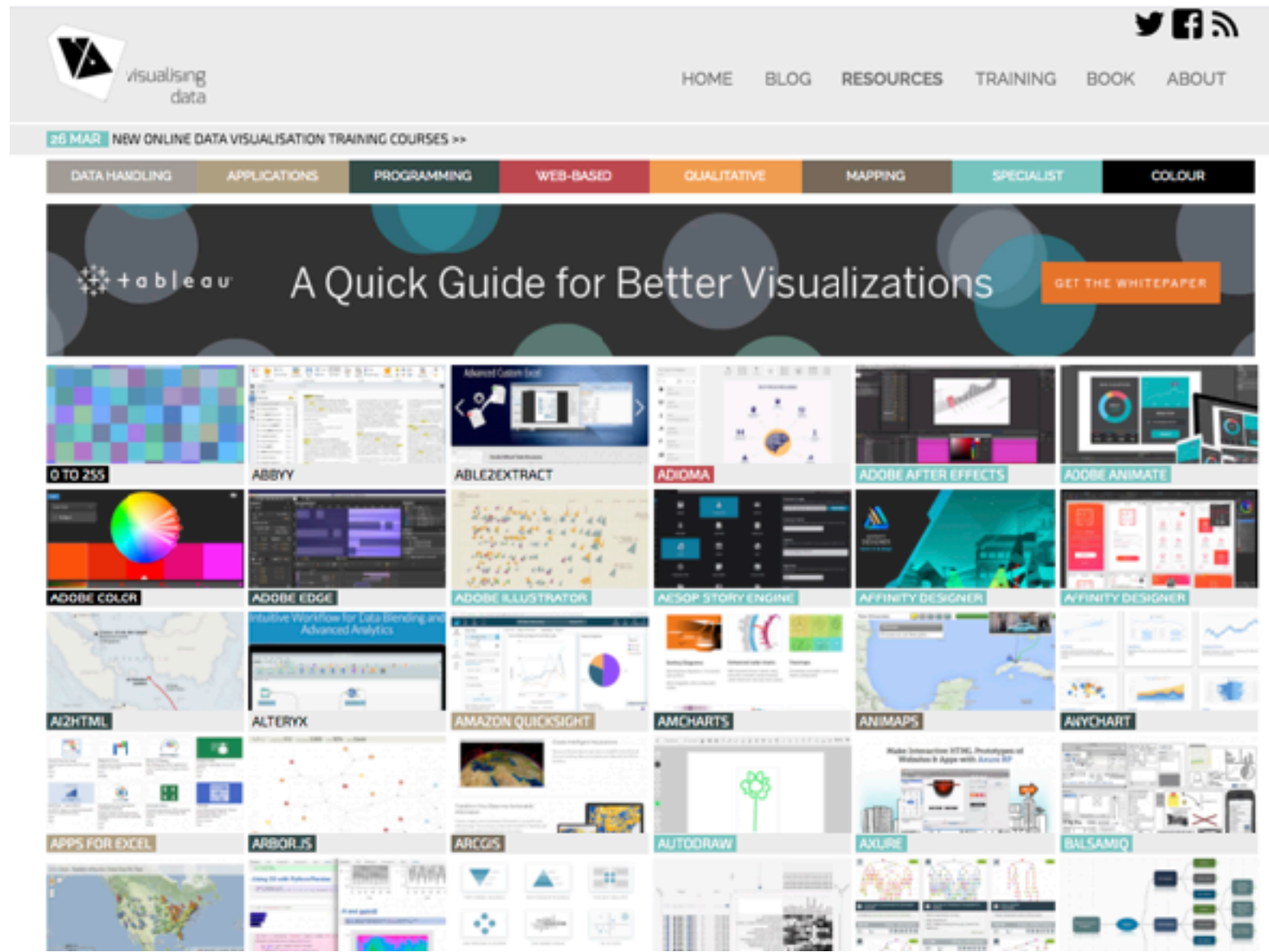
- pre-paper talk
 - write and give talk first, as if presenting at conference
 - iterate on talk slides to get structure, ordering, arguments right
 - then create paper outline from final draft of slides
 - encourages concise explanations of critical ideas, creation of key diagrams
 - avoids wordsmithing digressions and ratholes
 - easier to cut slides than prose you agonized over
- pre-paper/practice talk feedback session: at least 2-3x talk length
 - global comments, then slide by slide detailed discussion
 - nurture culture of internal critique (build your own critique group if necessary)
- have non-authors read paper before submitting
 - internal review can catch many problems
 - ideally group feedback session as above

Next Steps

Tools & ideas: Andy Kirk's Visualizing Data

<http://www.visualisingdata.com/resources/>

<https://www.visualisingdata.com/blog/>

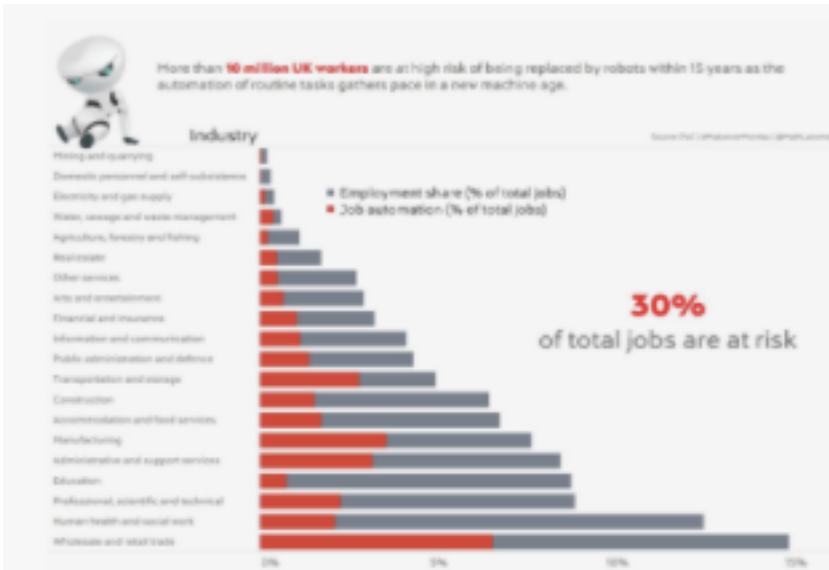


Videos

- many great conferences with free videos online
 - broadly accessible: OpenVisConf, Eyeo, InformationPlus
 - cutting-edge technical research: IEEE VIS

Redesign En Masse: **Makeover Mondays**

- easy entry point (Tableau focus)



Week 14 – Millions of UK workers at risk of being replaced by robots

Apr 7, 2017

During week 14 we looked at job automation and the potential impact of robots and AI on the UK employment market.



Week 13 – The Secret of Success

Mar 31, 2017

Week 13 took a look at a Russian survey about the secret of success. Dot plot, bump charts, bar charts, radar charts. This week had it all! Plus seven lessons to take on board.



Week 12 – March Madness

Mar 24, 2017

We looked at March Madness data for week 12, highlighting the phenomenon that is US college basketball. Quite a few vizzes showed the passion that

Visual Design Process In Depth: **Dear Data**

- inspiring celebration of data humanism

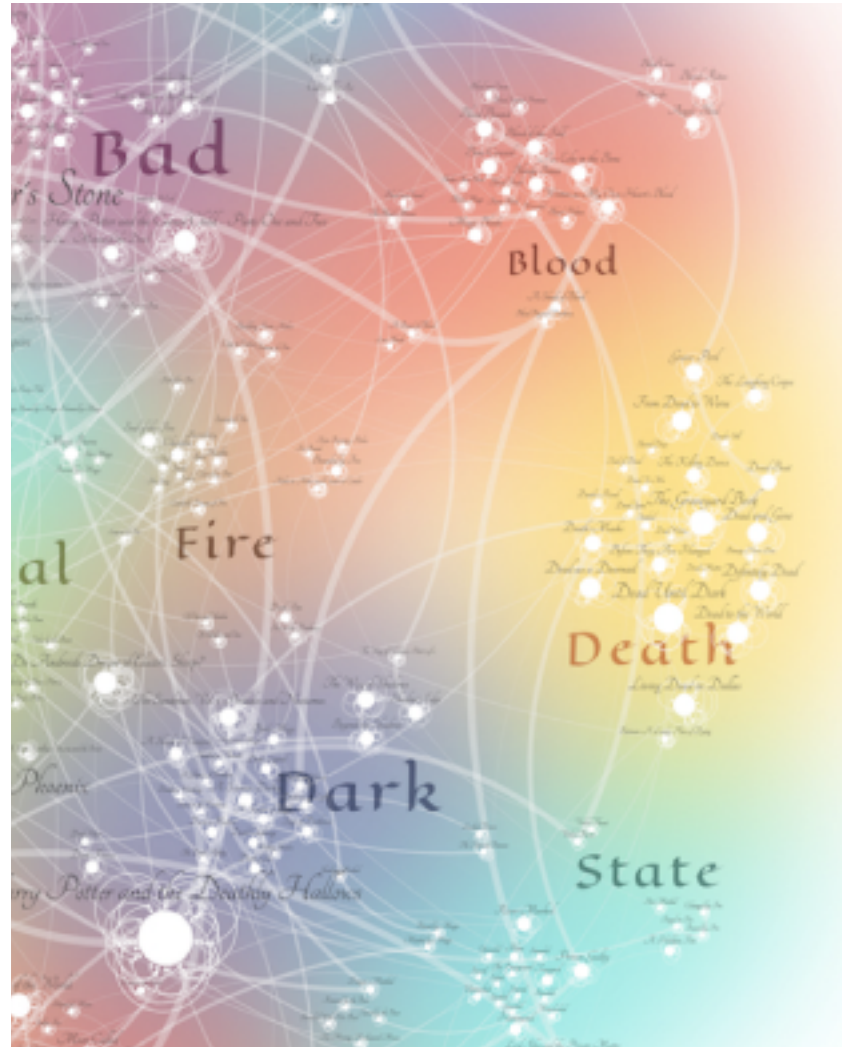


<http://www.dear-data.com/by-week/>

Giorgia Lupi and Stefanie Posavec

Visual Design Process In Depth: **Data Sketches**

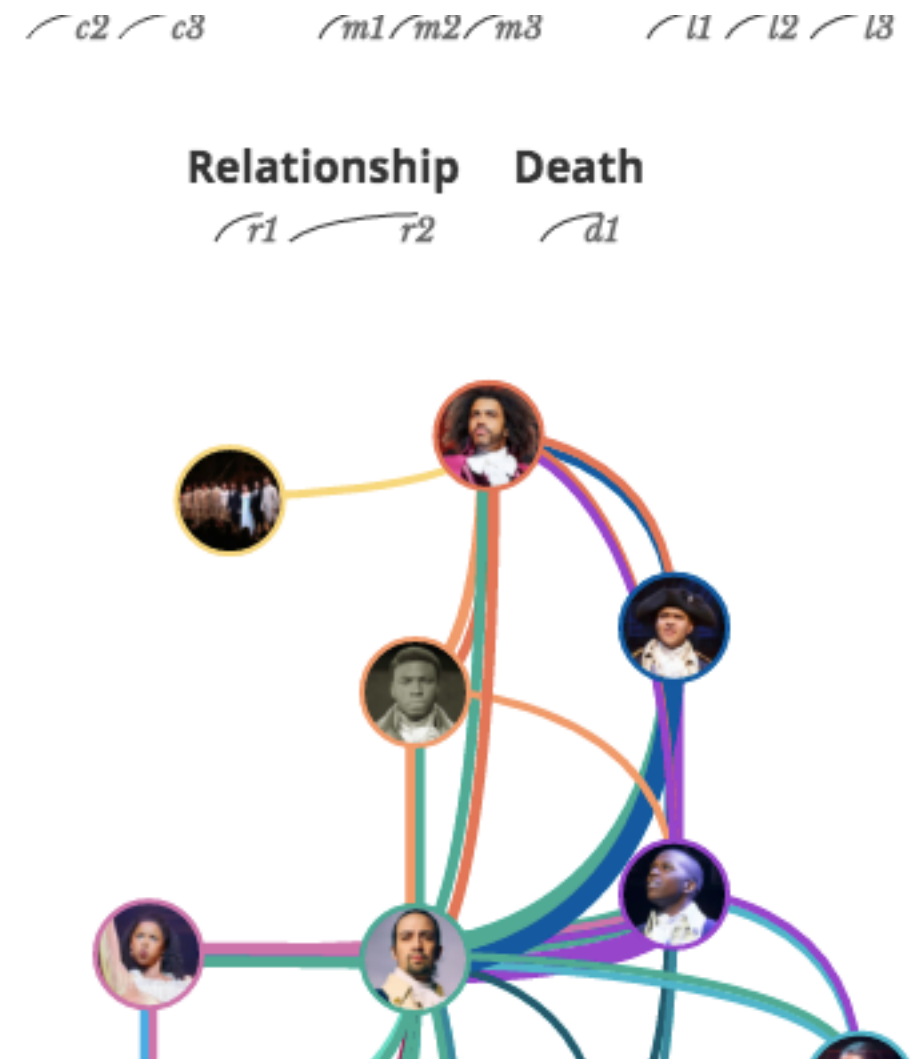
- detailed process notes, from sketching through coding



November
Books

Searching for patterns in Fantasy titles and musical lyrics

Read more...



<http://www.datasketch.es/>

Shirley Wu and Nadieh Brehmer

Pathways for more participation

- join Viz@UBC
 - <https://dfp.ubc.ca/initiatives/viz-ubc>
 - get on visatubc-announce email list (send mail to vizatubc-info@cs.ubc.ca)
 - talk series
- join Vancouver Visualization meetup
 - <https://www.meetup.com/Vancouver-Data-Visualization/>
 - 4K members
- join Data Visualization Society
 - <https://www.datavisualizationsociety.com/>
 - less than two years old, 10K+ members around the world
 - resources, jobs board, super-active Slack incl local groups, challenges, ...
 - Medium articles: [Nightingale](#)

Next Week

Come talk!

- encourage meeting with me to get advice/feedback before final present
 - chance to get feedback while you can still act on it
 - optional, not mandatory
 - do send email to schedule, can't meet with all 16 teams in last few days or in Tue office hours!