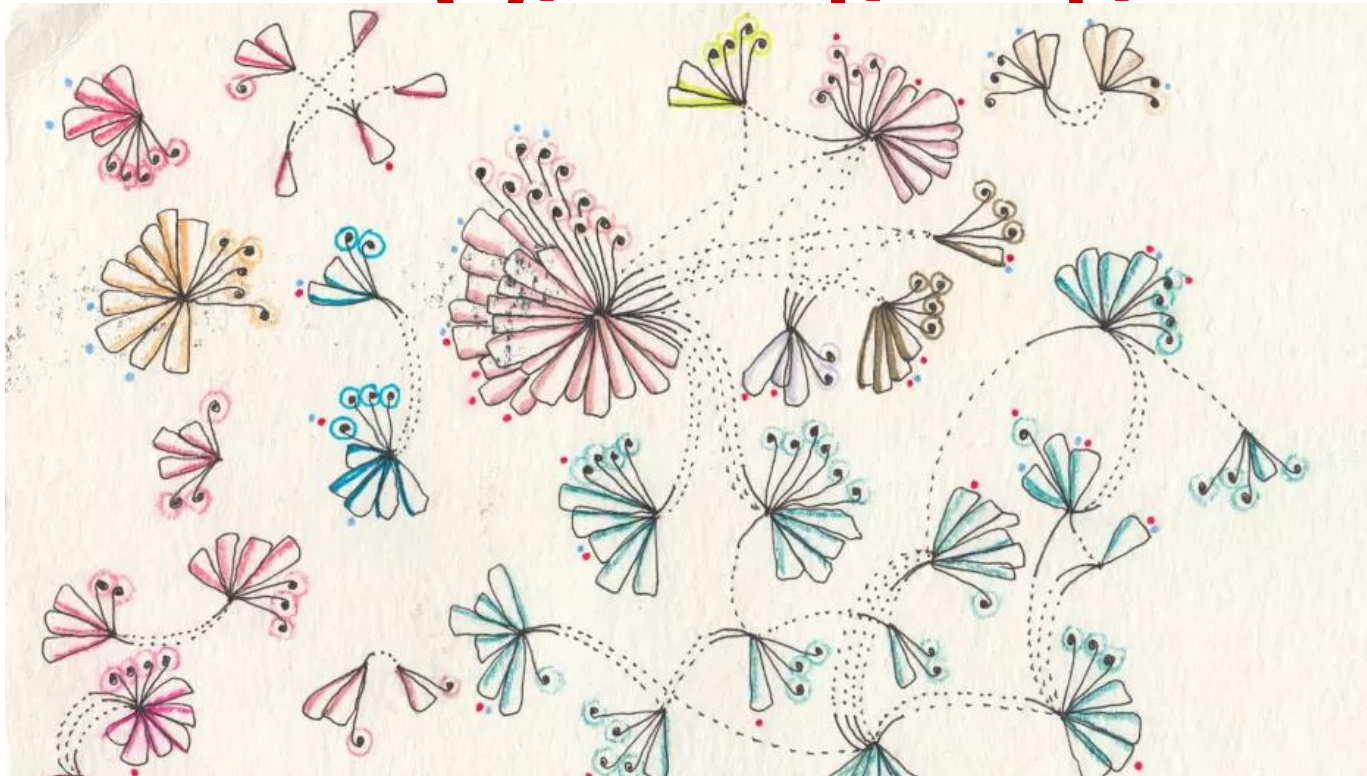


How to Design Beautiful



?

Maria Brbić
SNAP Workshop
Oct 19, 2021
(Adapted by Rex 2022)

Why Figures Matter?

- Opportunity to effectively convey complex information
- Well designed figures help audience understand your key ideas
- Often **the first part** seen by the editor/reviewers
 - Your figures can sell the paper!
 - You want to achieve the “WOW”



**A PICTURE IS
WORTH A
1,000 WORDS**

Today's Workshop

- Good practices
- How to design memorable figures?
- Guidelines and resources

Today's Workshop

- **Good practices**
- How to design memorable figures?
- Guidelines and resources

Rule #1: Know Your Audience

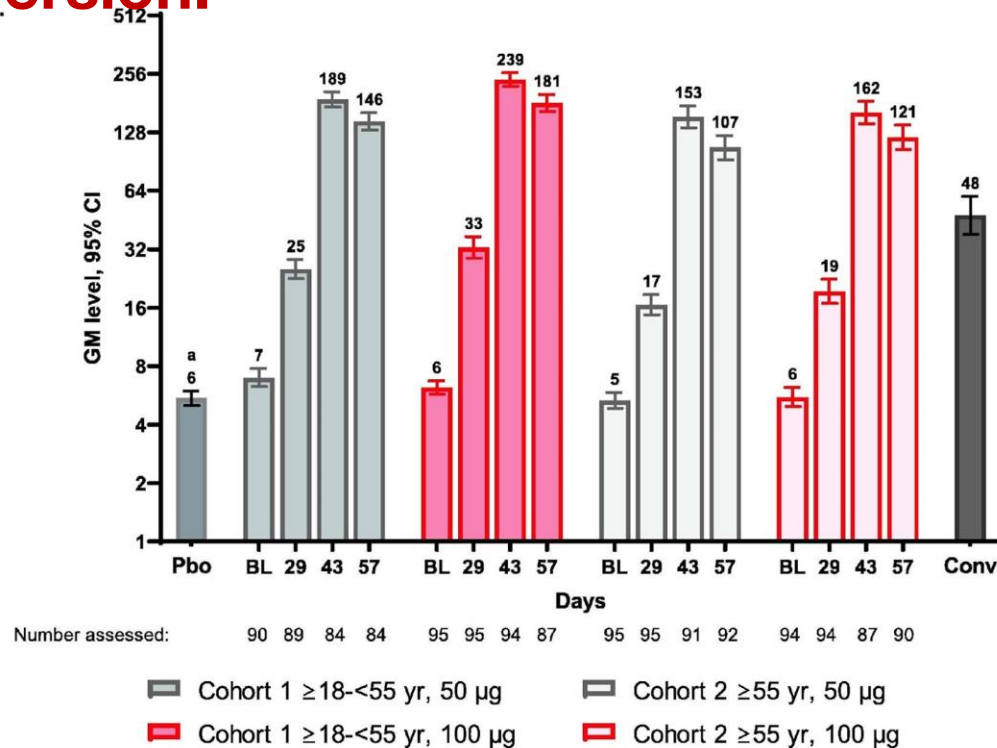
- ML conferences [NeurIPS, ICML, ICLR, ...]
 - You usually need 1 method overview figure + results figures
 - Audience familiar with key concepts
 - Follow graphical conventions and norms in the field
- Interdisciplinary journals [Nature, Science, PNAS,...]
 - Complex figures consisting of many panels
 - Broader audience [?] they may not know what you know!
 - Graphic elements need to be generally recognized and clearly identifiable

Rule #2: Less Is More

- Make simple & concise figures
- Focus on the key message you want to convey
- Avoid “chartjunk” [E. Tufte]
 - ❑ Unnecessary visual elements that do not improve the message
 - ❑ Are details clear to the general audience?
 - ❑ Do you need all the colors in your figure?
 - ❑ Do you need all the shapes in your figure?

Rule #2: Less Is More

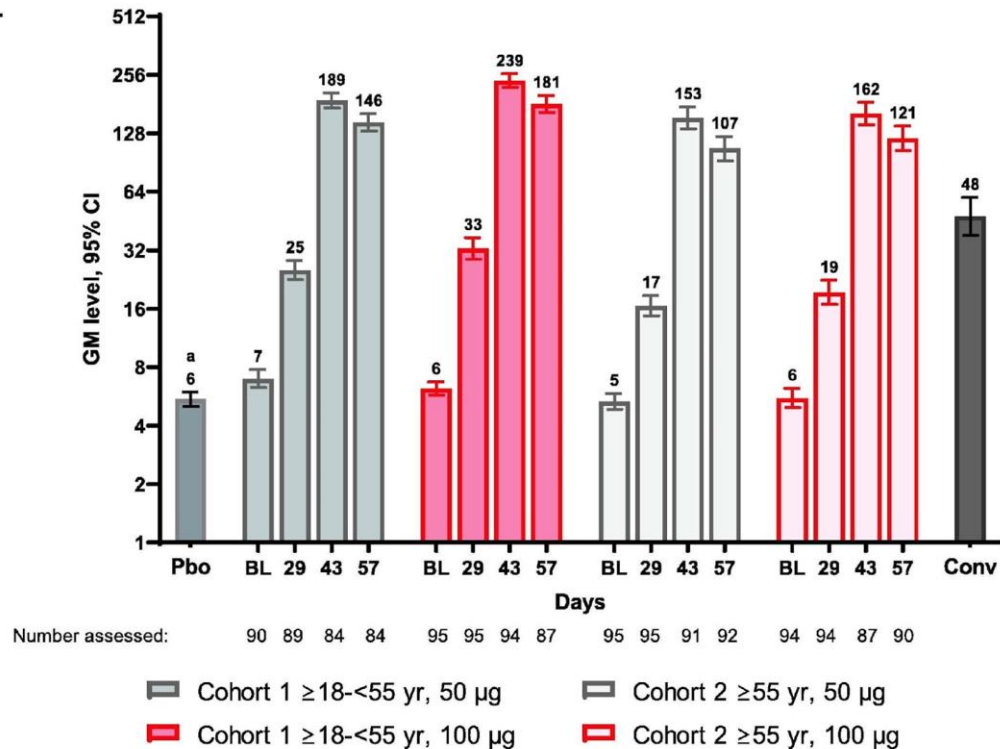
Published version:



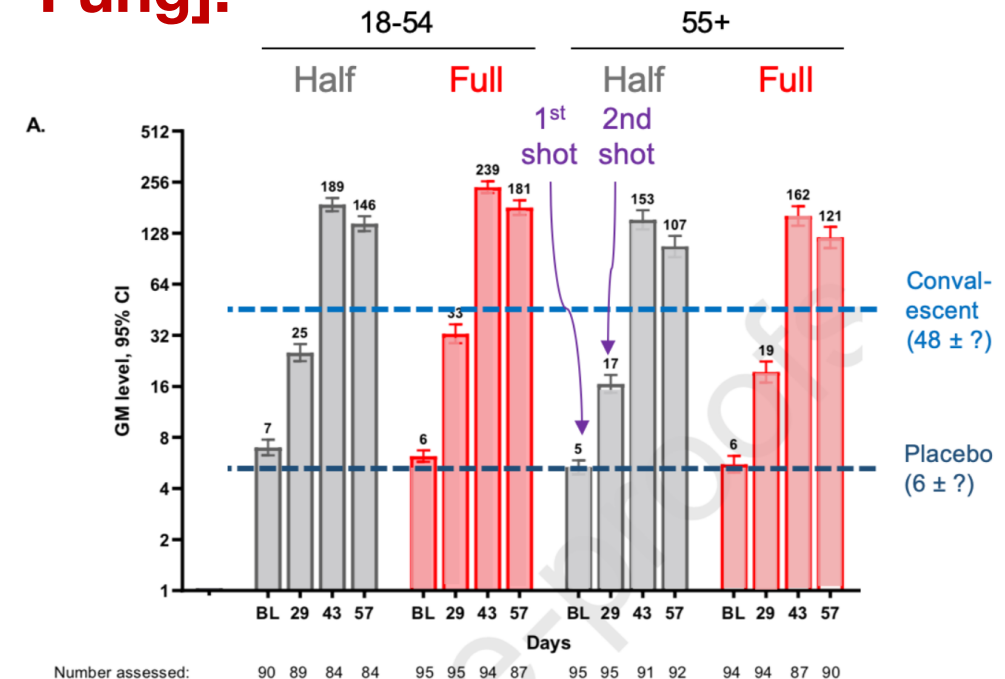
- The study's main goal is to compare half doses to full doses of the Moderna vaccine
- Age-group dimension is secondary – vaccine is expected to work better for younger people

Rule #2: Less Is More

Published version:



Improved version [K. Fung]:

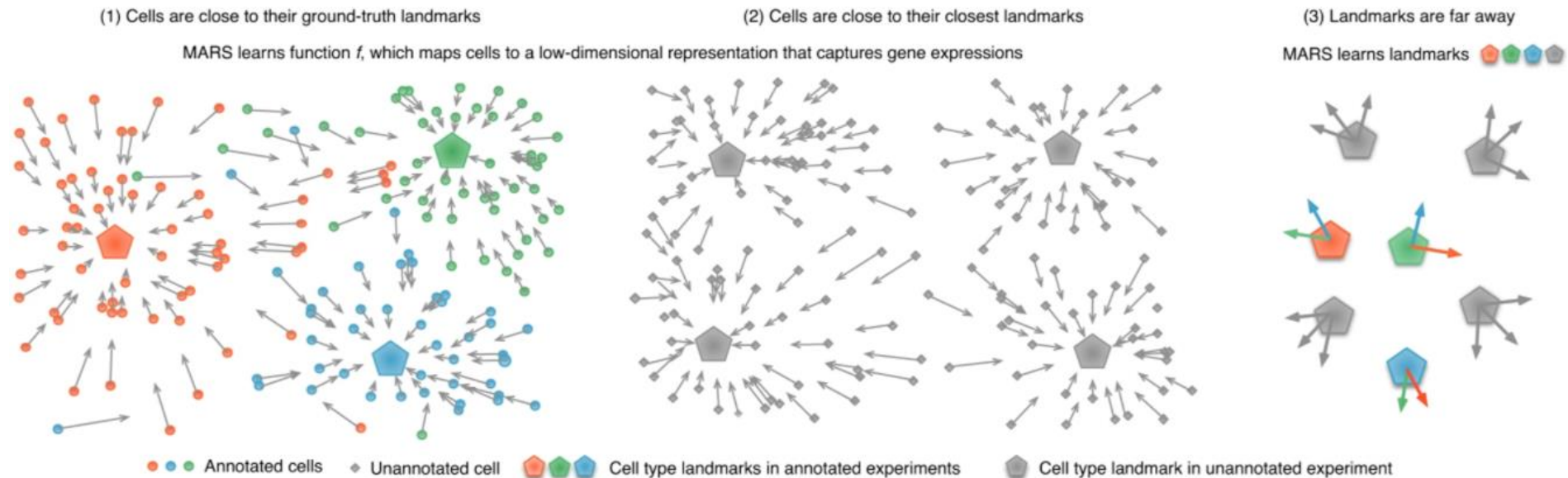


Adapted from Chu, et. al. 2021.

Kaiser Fung / JunkCharts

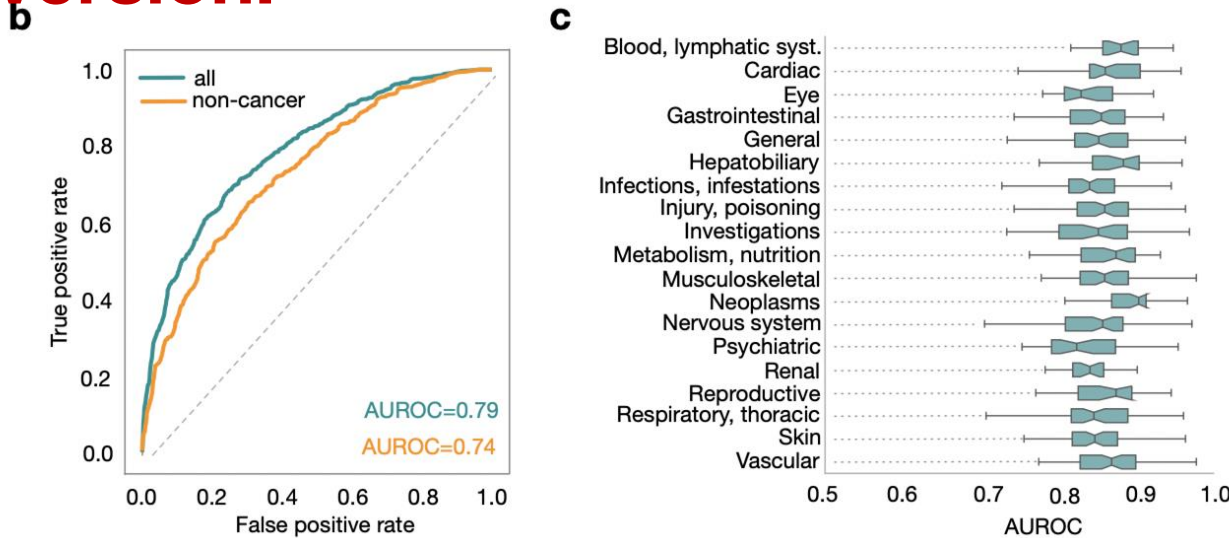
Rule #3: Annotate, annotate, annotate!

- **Annotate:**
 - Axes
 - Explain what elements represent
 - Ideally your figure is clear without reading the text

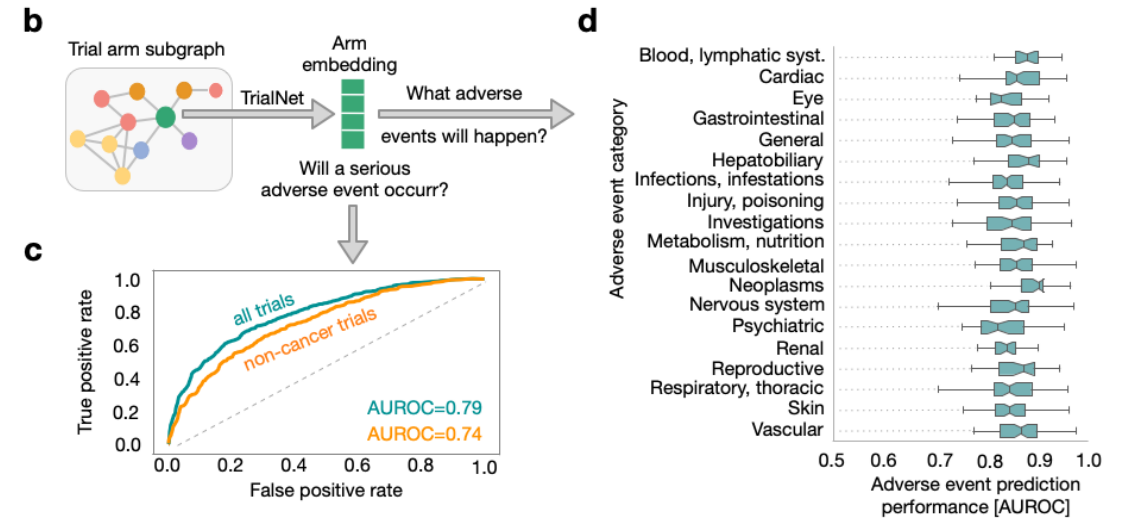


Rule #3: Annotate, annotate, annotate!

First version:



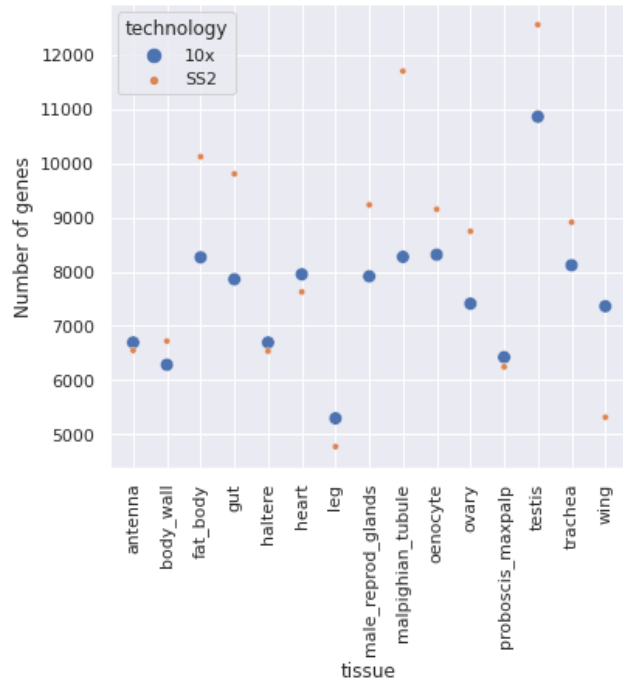
Improved version:



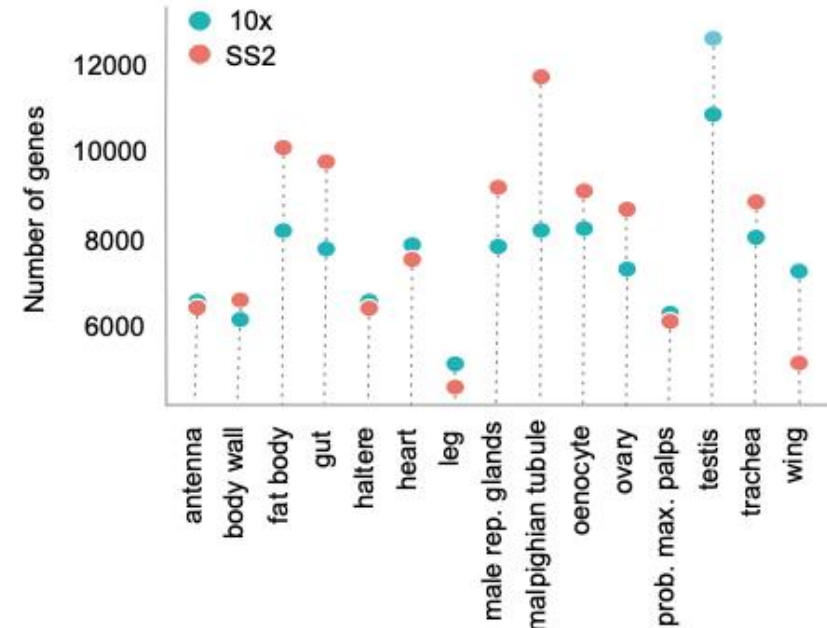
Rule #4: Never Trust Default Settings!

- **All plots** require at least some manual tuning
- Tune size, font, colors, styles, ticks, markers...

Default settings:



Tuned settings:



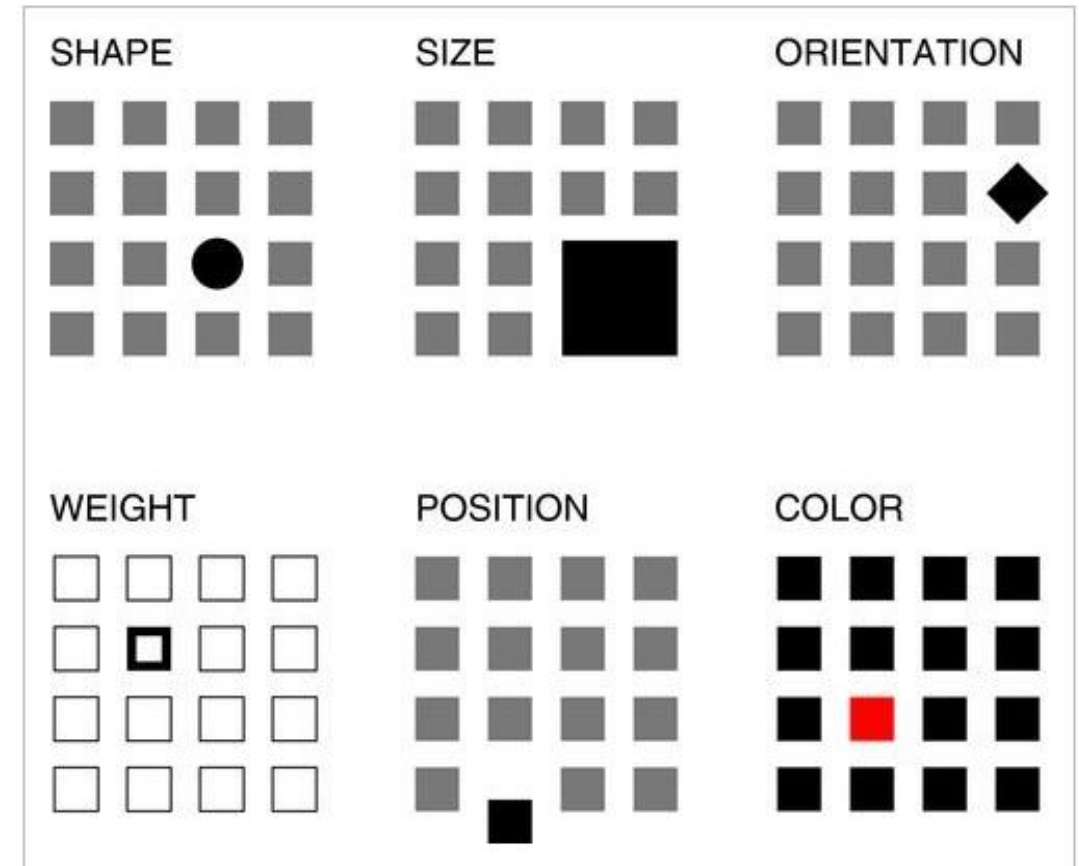
Rule #4: Never Trust Default Settings!

It is very common to adjust the figures through post-processing

1. Plot figures in Matplotlib / Seaborn
2. Adjust settings (axis font size, line width ...)
3. Save in a vector graphics format (.eps, .pdf)
4. **Open vector software (Adobe Illustrator, Inkscape, Sketch ...) and edit them!**

Rule #5: Use Visual Contrast

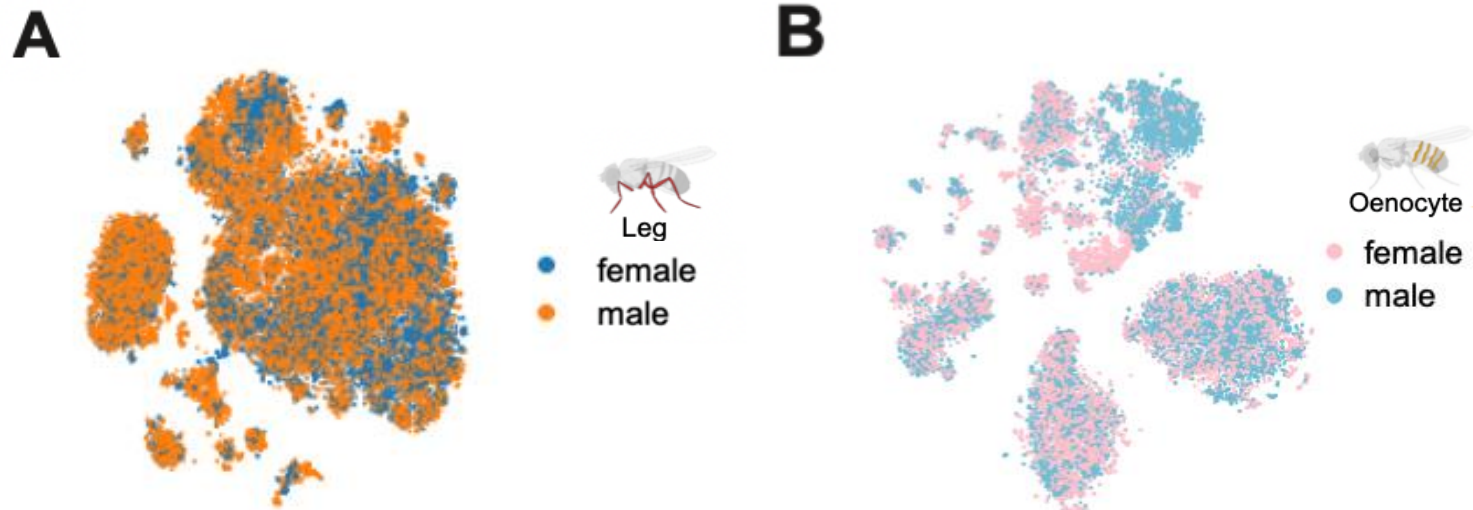
- Use visual contrast to highlight the important information
- Vary the size, shape, position, orientation, or color of an element to make the key part of the figure the most visually prominent
- Avoid incorporating every form of visual contrast



Rule #6: Be Consistent

- Follow your conventions
- Applies to colors, fonts, shapes, upper/lower case...

**Same
information
different color:
bad practice!**

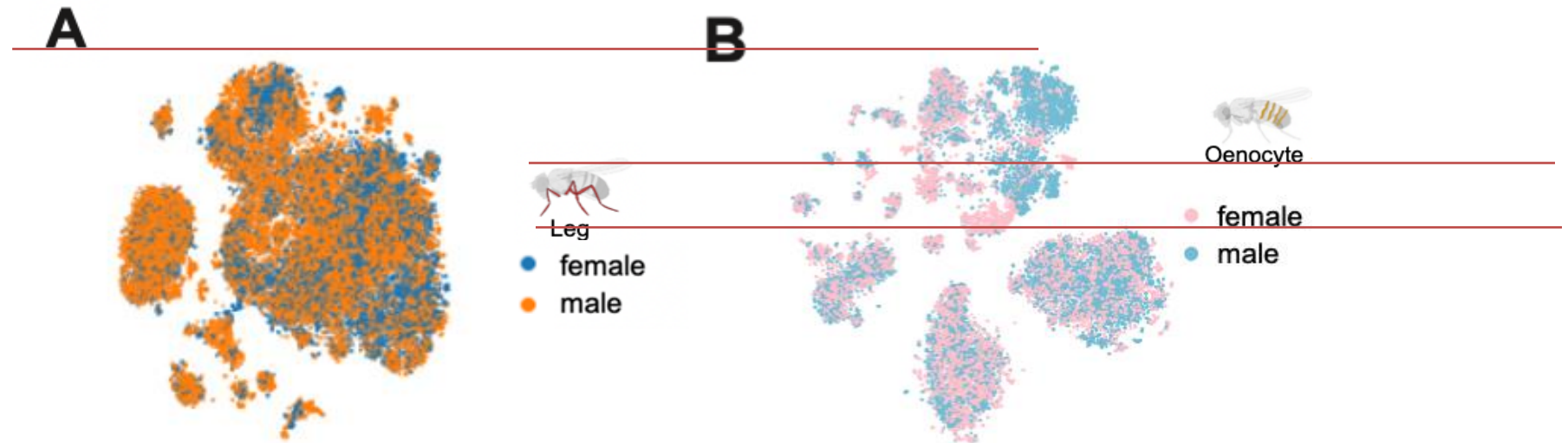


Rule #7: Check Alignment

- A grid can be a useful tool for ensuring alignment
- Every software has rulers

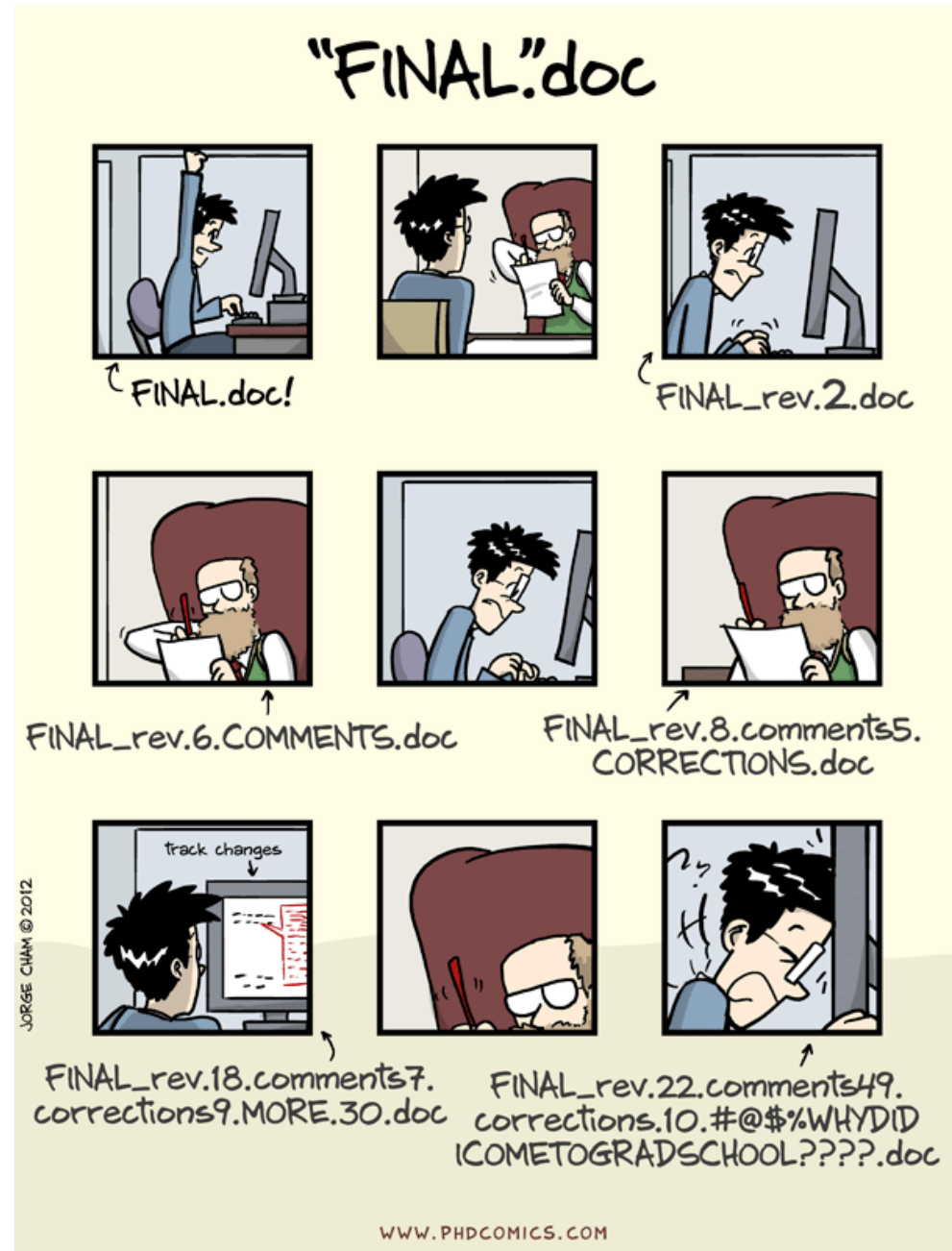


**Alignment is
not
appropriate:
bad practice!**



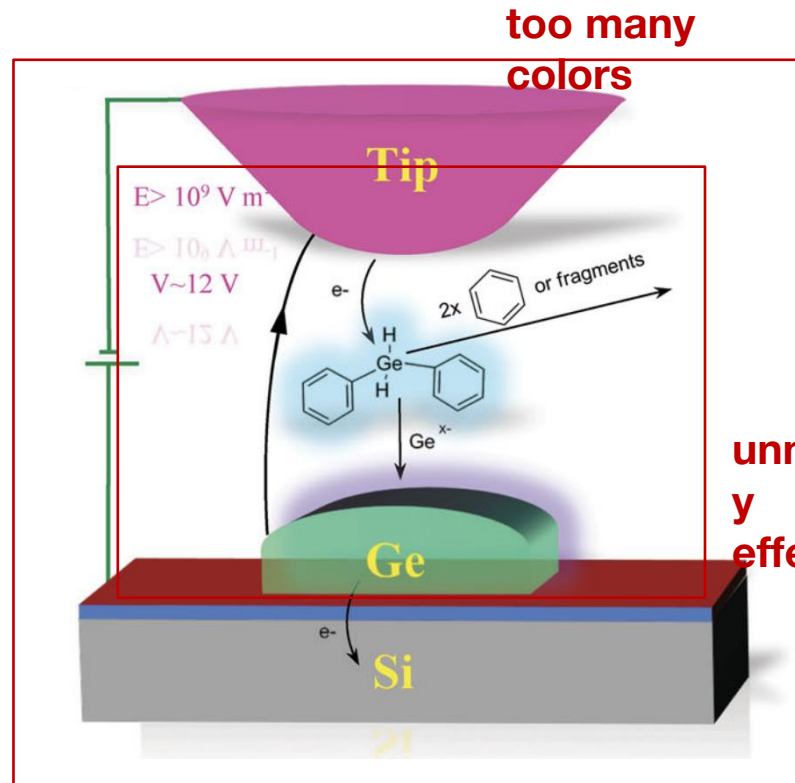
Rule #8: Figures Evolve!

- Your first version won't be good
- You usually need a number of iterations

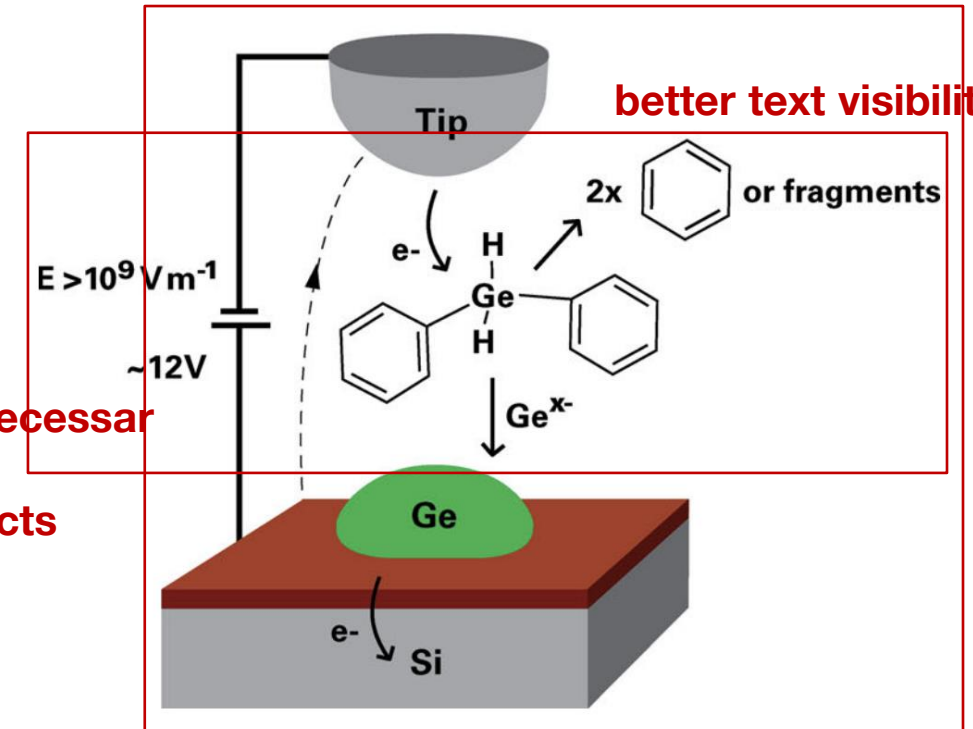


Rule #8: Figures Evolve!

Version 1



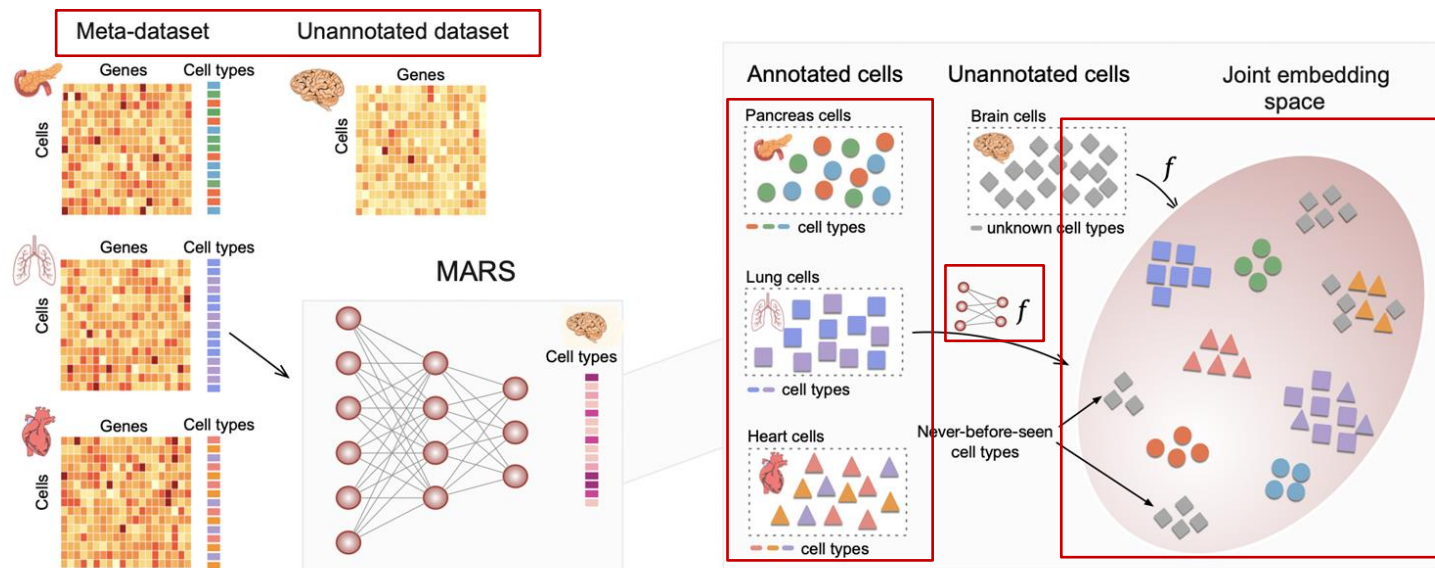
Version 2



better use of
components and
components are aligned
on an implied grid

Rule #8: Figures Evolve!

Version
1

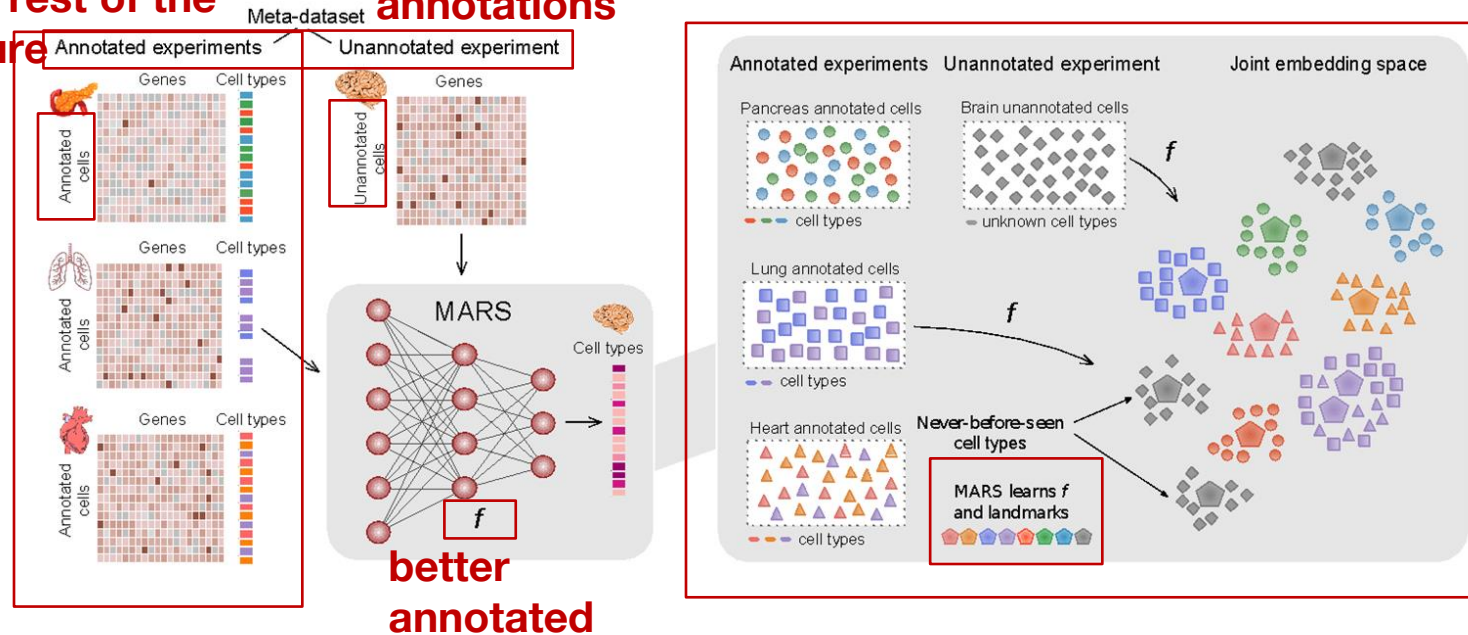


icons are
unnecessary detail
questionable radial blend
effect rounded corners

colors match
the rest of the
figure

consistent
annotations

Version
2



better
annotated

Today's Workshop

- Good practices
- **How to design memorable figures?**
- Guidelines and resources

Designing Memorable Figures: ML Conference

- Try to make it memorable!
- Visualize key insights of your approach
- Figure follows abstract
- Catchy name of the method

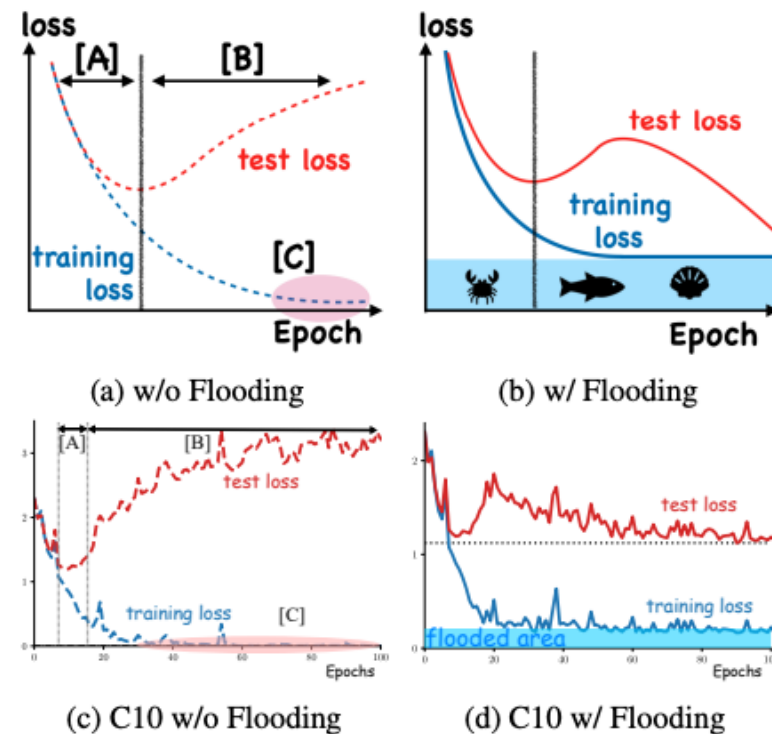
Approach is clear (& memorable) from the abstract and figure!

Do We Need Zero Training Loss After Achieving Zero Training Error?

Takashi Ishida^{1,2} Ikko Yamane¹ Tomoya Sakai³ Gang Niu² Masashi Sugiyama^{2,1}

Abstract

Overparameterized deep networks have the capacity to memorize training data with zero *training error*. Even after memorization, the *training loss* continues to approach zero, making the model overconfident and the test performance degraded. Since existing regularizers do not directly aim to avoid zero training loss, it is hard to tune their hyperparameters in order to maintain a fixed/preset level of training loss. We propose a direct solution called *flooding* that intentionally prevents further reduction of the training loss when it reaches a reasonably small value, which we call the *flood level*. Our approach makes the loss float around the flood level by doing mini-batched gradient descent as usual but gradient ascent if the training loss is below the flood level. This can be implemented with one line of code and is compatible with any stochastic optimizer and other regularizers. With flooding, the model will continue to “random walk” with the same non-zero training loss, and we expect it to drift into an area with a flat loss landscape that leads to better generalization. We experimentally show that flooding improves performance and, as a byproduct, induces a double descent curve of the test loss.



Designing Memorable Figures: Journals

- Try to make it memorable!
- Visualize key insights of your approach
- Figure follows abstract
- Catchy name of the method

Approach is clear (& memorable) from the abstract and figure!

Efficient integration of heterogeneous single-cell transcriptomes using Scanorama

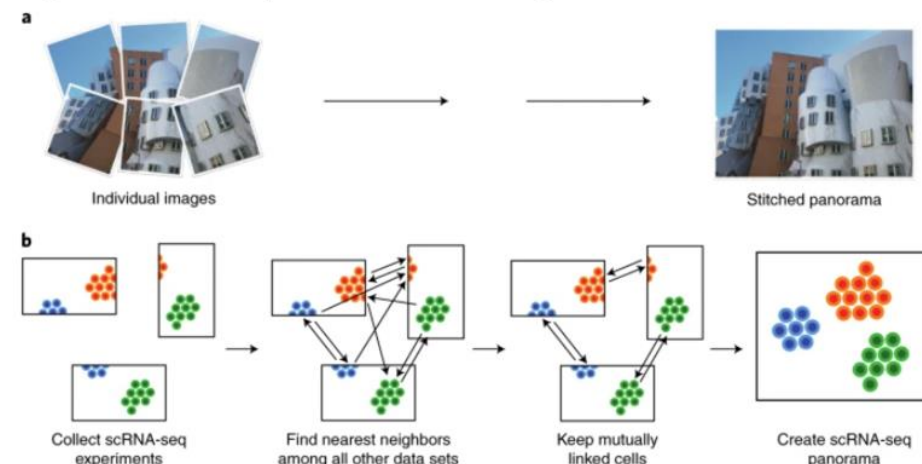
Brian Hie, Bryan Bryson  & Bonnie Berger 

Nature Biotechnology **37**, 685–691 (2019) | [Cite this article](#)

Abstract

Integration of single-cell RNA sequencing (scRNA-seq) data from multiple experiments, laboratories and technologies can uncover biological insights, but current methods for scRNA-seq data integration are limited by a requirement for datasets to derive from functionally similar cells. We present Scanorama, an algorithm that identifies and merges the shared cell types among all pairs of datasets and accurately integrates heterogeneous collections of scRNA-seq data. We applied Scanorama to integrate and remove batch effects across 105,476 cells from 26 diverse scRNA-seq experiments representing 9 different

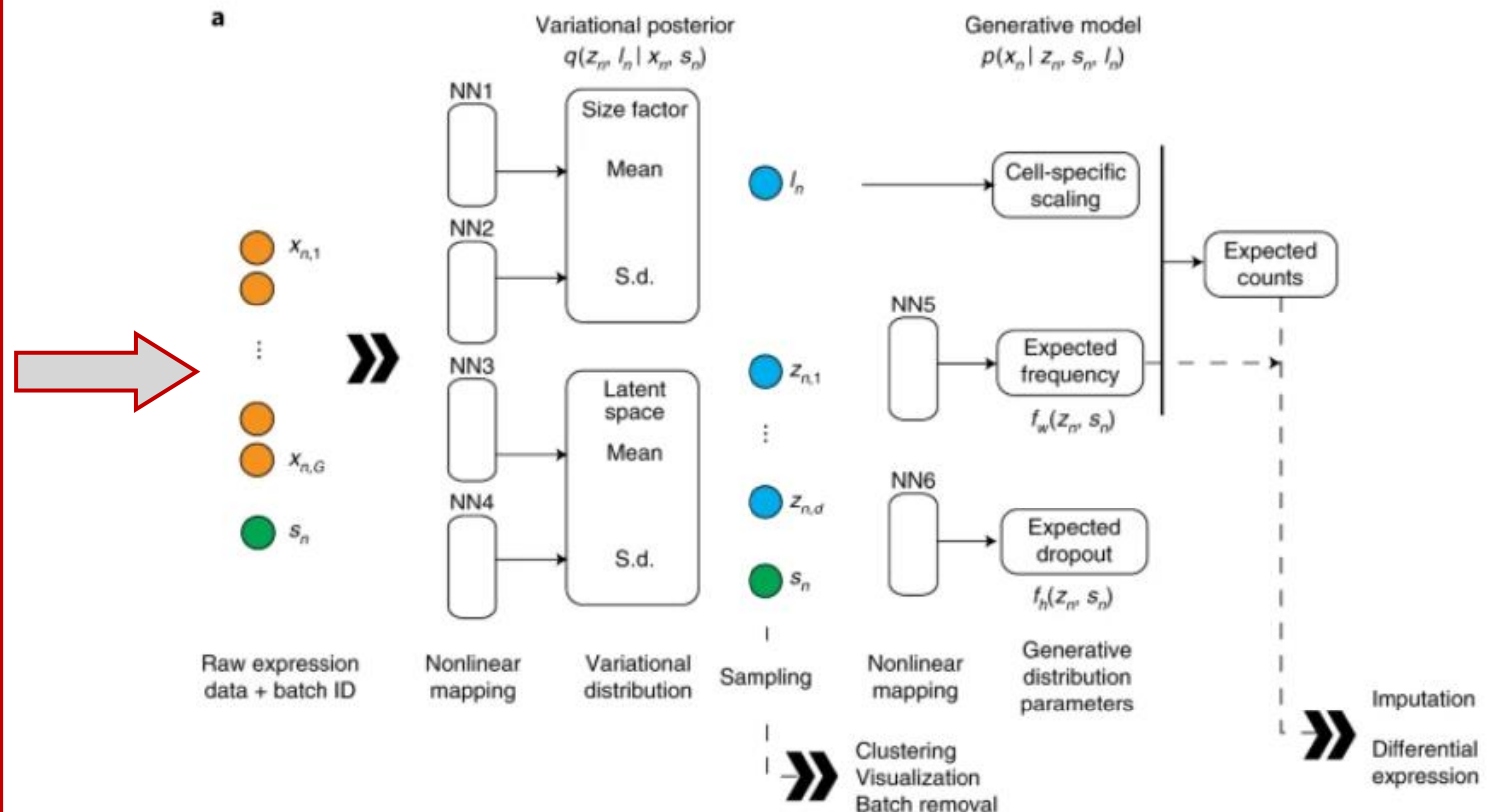
Fig. 1: Illustration of ‘panoramic’ dataset integration.



Designing Memorable Figures: Journals

- Key insight/idea missing: What is unique about this approach? What is this method doing?
- Too many details
- Many symbols that are not introduced
- ...
- Nothing memorable about it

What you do not want to do:



Brainstorm

- Try to think about a paper for which you can remember the figure very well (ideally paper you've seen only once)
 - This is how you want to design your figures!
- Now try to think about the paper you know very well but you can't remember the figure
 - This is how you do not want to design your figures!
- Go back to these papers and find out why they are (not) memorable!

Today's Workshop

- Good practices
- How to design memorable figures?
- **Guidelines and resources**

Steps to Follow

1. Make a **plot**
 - Using Python, R, MATLAB,...
 - For network data Cytoscape, Gephi...
2. Export figure to a **vector format file**
 - PDF, EPS, SVG ...
 - Vector graphics images are editable & can be resized without loss of quality!
3. Open figure in **graphics editor**
 - OmniGraffle, Adobe Illustrator, Inkscape, GIMP,...
4. Connect panels, annotate, improve

Visualization Libraries & Tools

- Python libraries: [Matplotlib](#), [ggplot](#), [Pandas](#), [Seaborn](#), [plotly](#), [Geoplotlib](#), [Bokeh](#)
- Latex: [TikZ](#), [PSTricks](#)
- Biorender: <https://biorender.com/>
- Network data: [Gephi](#), [Cytoscape](#), [hive plots](#), [igraph](#), [NetworkX](#), [Graphviz](#)
- Javascript: [D3.js](#)
- Colors selections: [RapidTables](#), [ColorBrewer](#)

Links & References

- E. Tufte: [The Visual Display of Quantitative Information](#)
- Rougier, Droettboom, Bourne: [Ten Simple Rules for Better Figures](#), Plos Comp Bio '14
- Rolandi, Cheng, Perez-Kriz: [A Brief Guide to Designing Effective Figures for the Scientific Paper](#), Advanced Materials '11
- Kaiser Fung Junk charts blog: https://junkcharts.typepad.com/junk_charts/
- Plotly blog: <https://blog.plotly.com/>
- DataViz Weekly overview blog: <https://www.anychart.com/blog/>