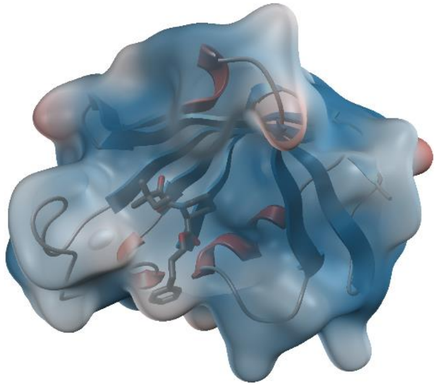


EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Visualization of Biological Data

Winter Term 2018/2019

Jun.-Prof. Dr. Michael Krone

Organization

- Lecturer
  - Jun.-Prof. Dr. Michael Krone (C308b)
  - michael.krone@uni-tuebingen.de
- Tutors
  - Mirjam Figaschewski (C305)
    e-mail: figasch@informatik.uni-tuebingen.de
  - Marco Schäfer (C308b)
    e-mail: t.b.a.
- Course materials, assignments, announcements etc. via ILIAS
  - Course title: **BIO4364 - Visualization of Biological Data - WS2018/19**
  - Registration with course password: **VisBio18**
  - *Please check ILIAS page regularly for announcements!*

Tutorials, Assignments, and Evaluation

- Tutorials (problem sessions)
  - Time slot: Monday, 12:00-14:00 (c.t.), room C215 → **starts next week!**
  - Short outline by **Mirjam** or **Marco**
  - Discussion of last week's assignment

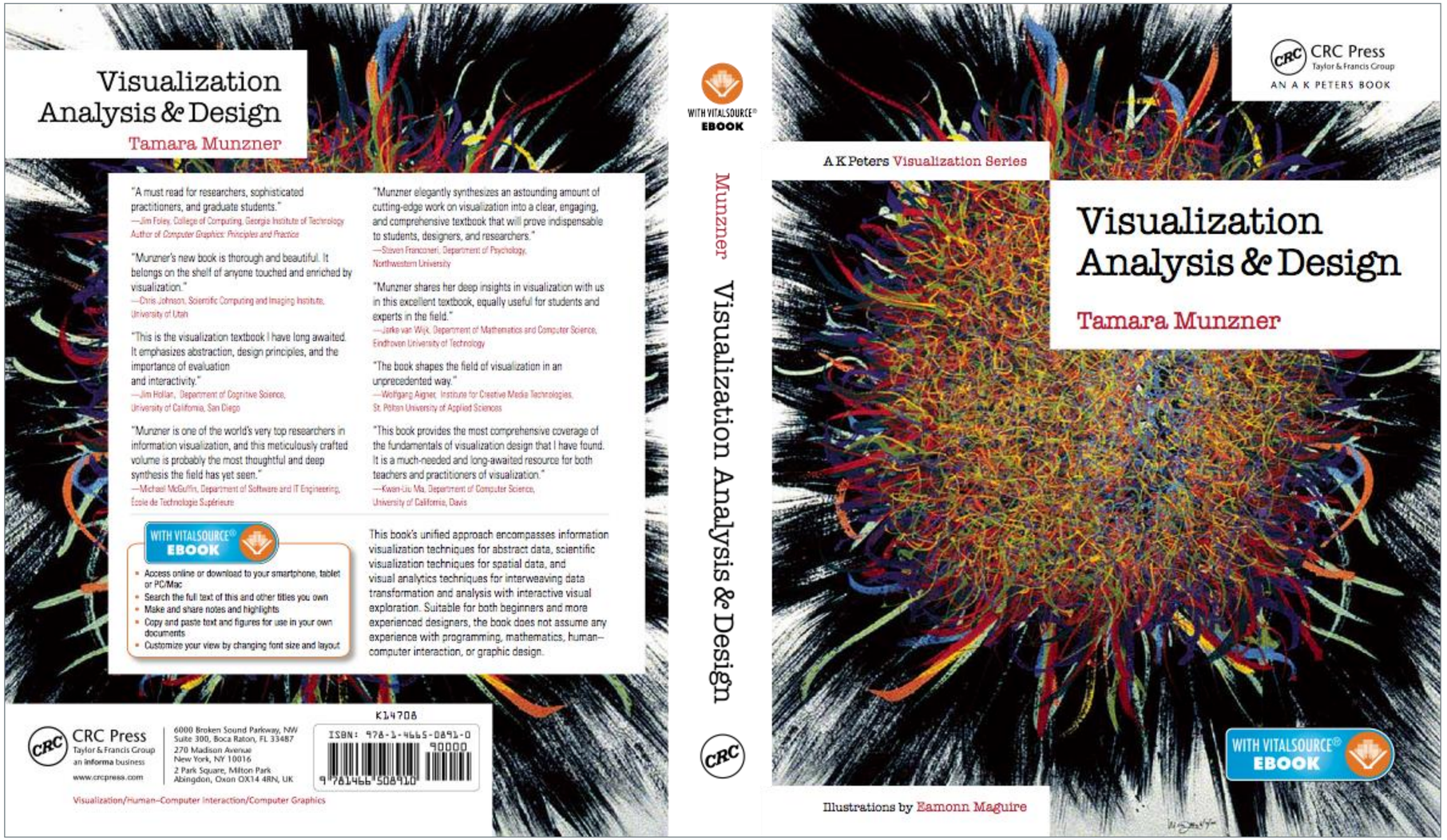
  → *Time for questions (about assignments and lecture – in addition to lecture!)*
- Assignments
  - 6 weekly assignments (until Christmas)
  - Rest of the semester: team project (~4-5 people)
- Requirement to take exam:
  - Pass assignments (>50% of points)
  - Possibility of bonus points for the exam
- Final exam determines grade

Lecture Content

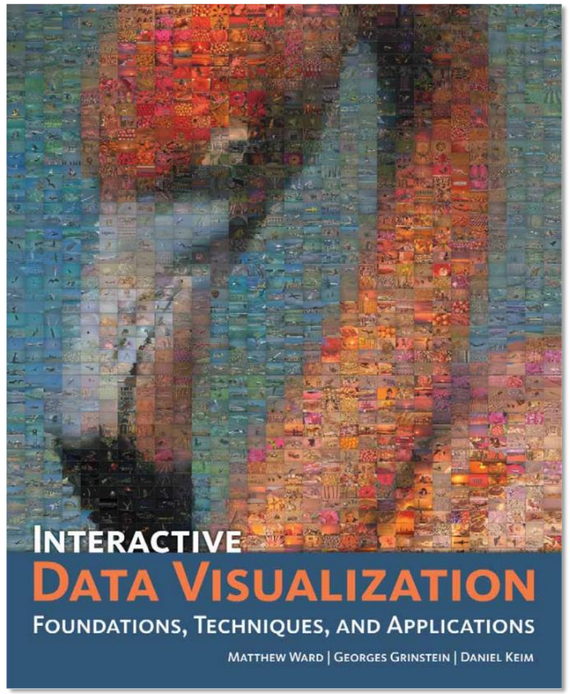
- Fundamentals of data visualization
- Applications for biological data
- Case studies

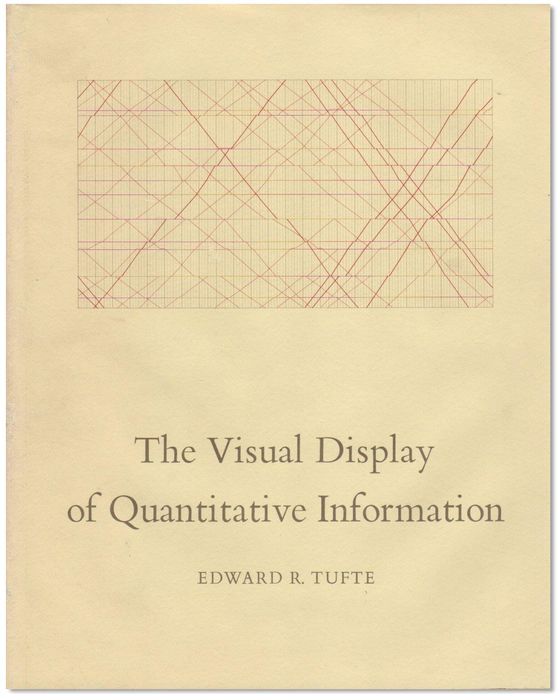
- Slides based on previous lectures by
  - Monika Balvočiūtė
  - Anna Górska
  - Apl. Prof. Dr. Kay Nieselt (C326a)
    - kay.nieselt@uni-tuebingen.de
  - Jun.-Prof. Dr. Julian Heinrich

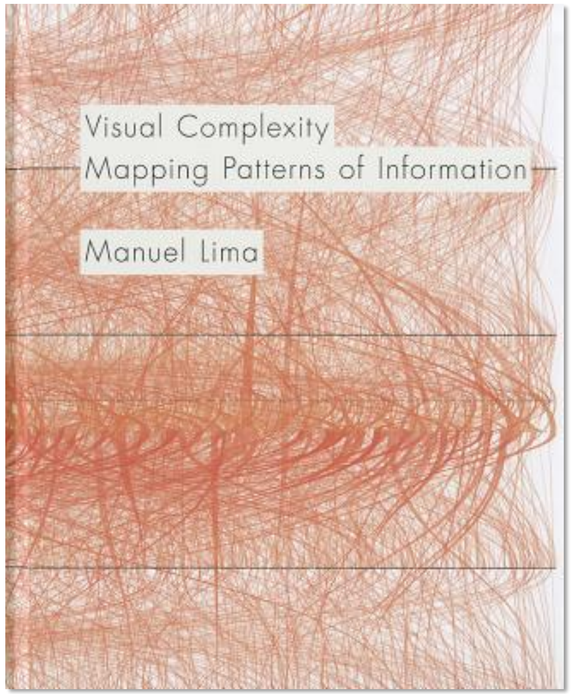
Book

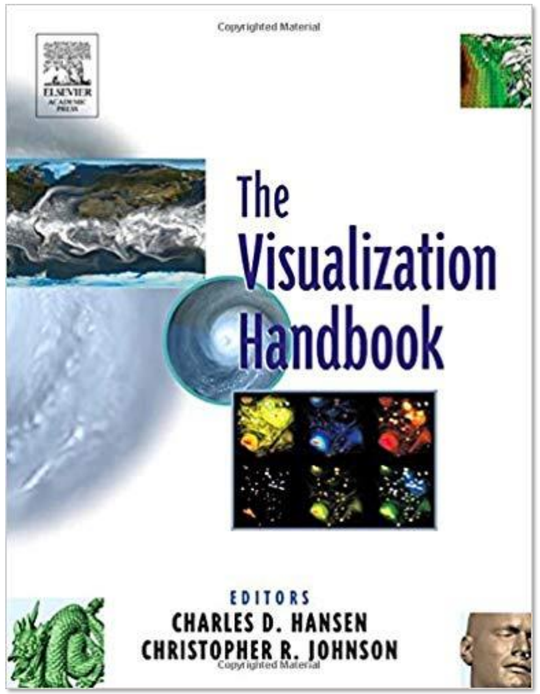

Visualization
Analysis & Design
Tamara Munzner
"A must read for researchers, sophisticated practitioners, and graduate students."
—Jim Foley, College of Computing, Georgia Institute of Technology
Author of Computer Graphics: Principles and Practice
"Munzner's new book is thorough and beautiful. It belongs on the shelf of anyone touched and enriched by visualization."
—Chris Johnson, Scientific Computing and Imaging Institute, University of Utah
"This is the visualization textbook I have long awaited. It emphasizes abstraction, design principles, and the importance of evaluation and interactivity."
—Jim Hollan, Department of Cognitive Science, University of California, San Diego
"Munzner is one of the world's very top researchers in information visualization, and this meticulously crafted volume is probably the most thoughtful and deep synthesis the field has yet seen."
—Michael McGuffin, Department of Software and IT Engineering, École de Technologie Supérieure
"Munzner elegantly synthesizes an astounding amount of cutting-edge work on visualization into a clear, engaging, and comprehensive textbook that will prove indispensable to students, designers, and researchers."
—Steven Franconeri, Department of Psychology, Northwestern University
"Munzner shares her deep insights in visualization with us in this excellent textbook, equally useful for students and experts in the field."
—Jarke van Wijk, Department of Mathematics and Computer Science, Eindhoven University of Technology
"The book shapes the field of visualization in an unprecedented way."
—Wolfgang Aigner, Institute for Creative Media Technologies, St. Pölten University of Applied Sciences
"This book provides the most comprehensive coverage of the fundamentals of visualization design that I have found. It is a much-needed and long-awaited resource for both teachers and practitioners of visualization."
—Kwan-Liu Ma, Department of Computer Science, University of California, Davis
WITH VITALSOURCE® EBOOK
Access online or download to your smartphone, tablet or PC/Mac
Search the full text of this and other titles you own
Make and share notes and highlights
Copy and paste text and figures for use in your own documents
Customize your view by changing font size and layout
This book's unified approach encompasses information visualization techniques for abstract data, scientific visualization techniques for spatial data, and visual analytics techniques for interweaving data transformation and analysis with interactive visual exploration. Suitable for both beginners and more experienced designers, the book does not assume any experience with programming, mathematics, human–computer interaction, or graphic design.
K14708
CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com
6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK
ISBN: 978-1-4665-0891-0
90000
9 781466 508910
Visualization/Human–Computer Interaction/Computer Graphics
WITH VITALSOURCE® EBOOK
Munzner
Visualization Analysis & Design
CRC Press
Taylor & Francis Group
AN A K PETERS BOOK
A K Peters Visualization Series
Visualization
Analysis & Design
Tamara Munzner
WITH VITALSOURCE® EBOOK
Illustrations by Eamonn Maguire

More Books…


INTERACTIVE
DATA VISUALIZATION
FOUNDATIONS, TECHNIQUES, AND APPLICATIONS
MATTHEW WARD | GEORGES GRINSTEIN | DANIEL KEIM

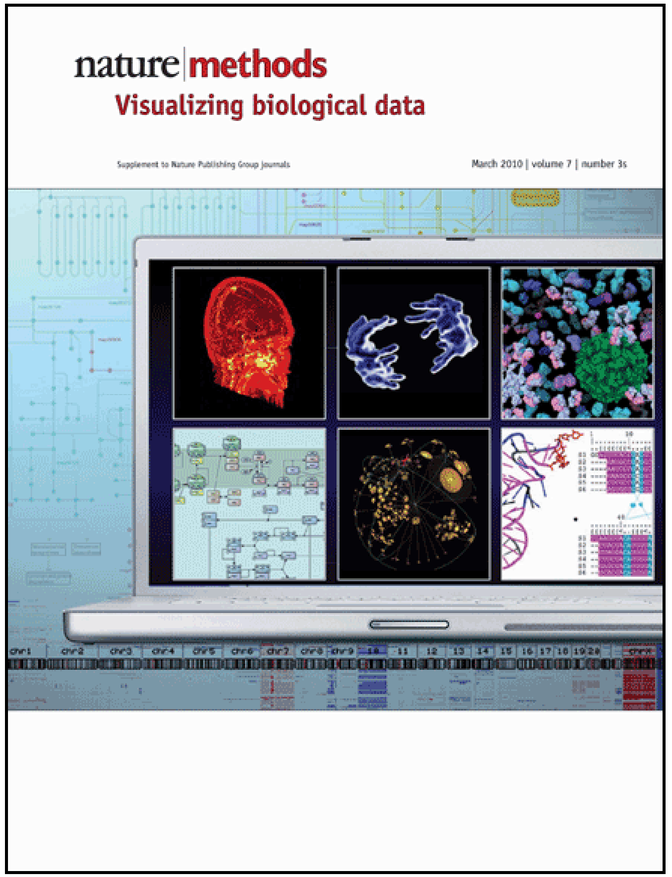

The Visual Display
of Quantitative Information
EDWARD R. TUFTE


Visual Complexity
Mapping Patterns of Information
Manuel Lima


Copyrighted Material
ELSEVIER
The Visualization Handbook
EDITORS
CHARLES D. HANSEN
CHRISTOPHER R. JOHNSON
Copyrighted Material

Other resources

- Nature methods special issue on biological visualization
- Points of view column
- Other research papers
  - Search online, e.g.:
    PubMed, IEEE Xplore Digital Library, Google Scholar…



Organization – Final Remarks

- **Asking questions during the lecture is highly encouraged!**
- I will try to start with some questions recapitulating the last lecture
  → Only works with your active participation!
- Further possibilities for questions:
  - I have no fixed office hours, but you can always arrange a meeting via e-mail…
    - …or just approach me after the lecture
  - Tutors (Mirjam, Marco)
  - Other students (e.g., via ILIAS Forum)
    → *Please don't ask for assignment solutions (or post them!)*

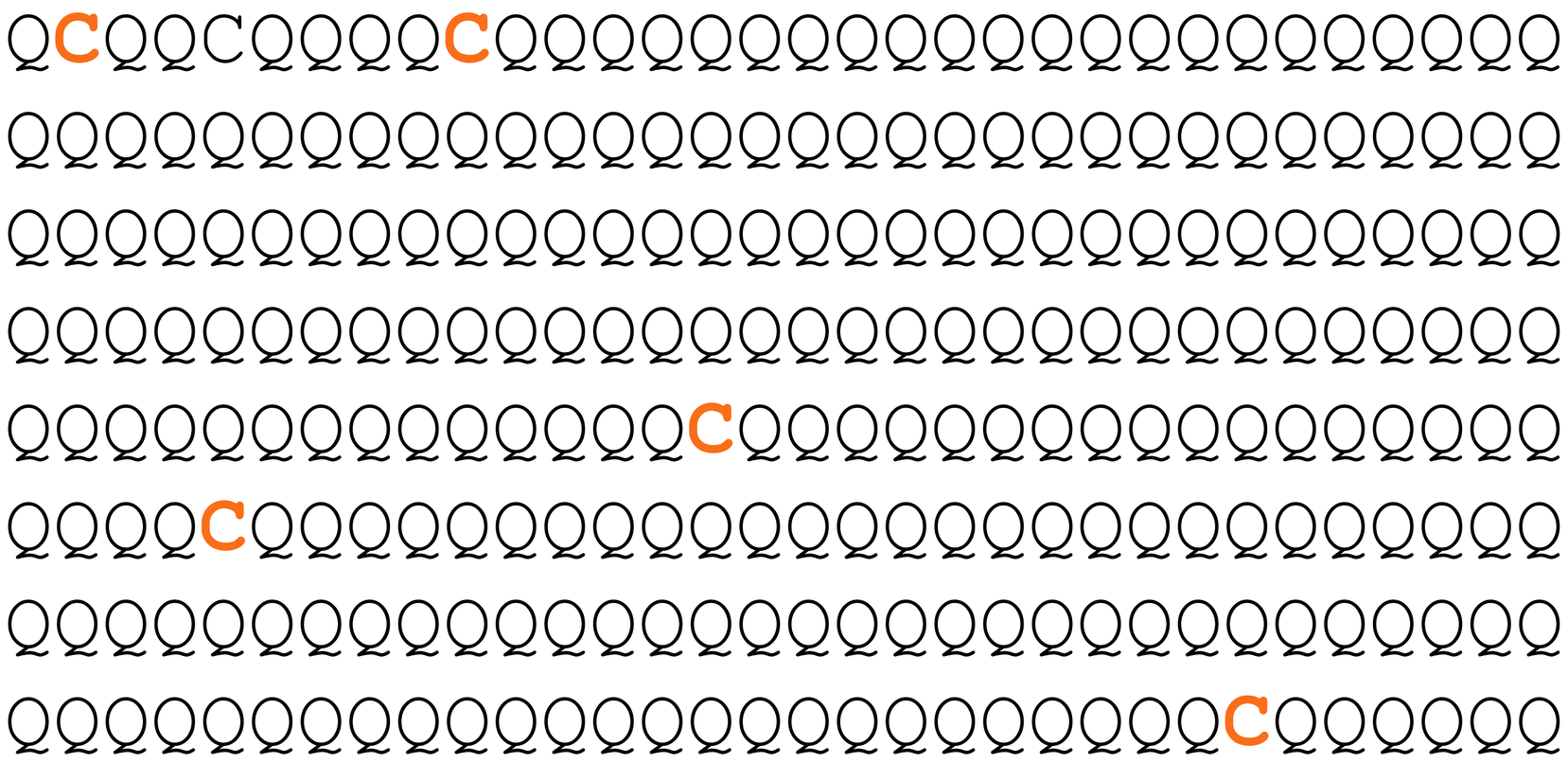
- **Short break after 45 min of each lecture?? (show of hands, please!)**

Why Use Visualization?

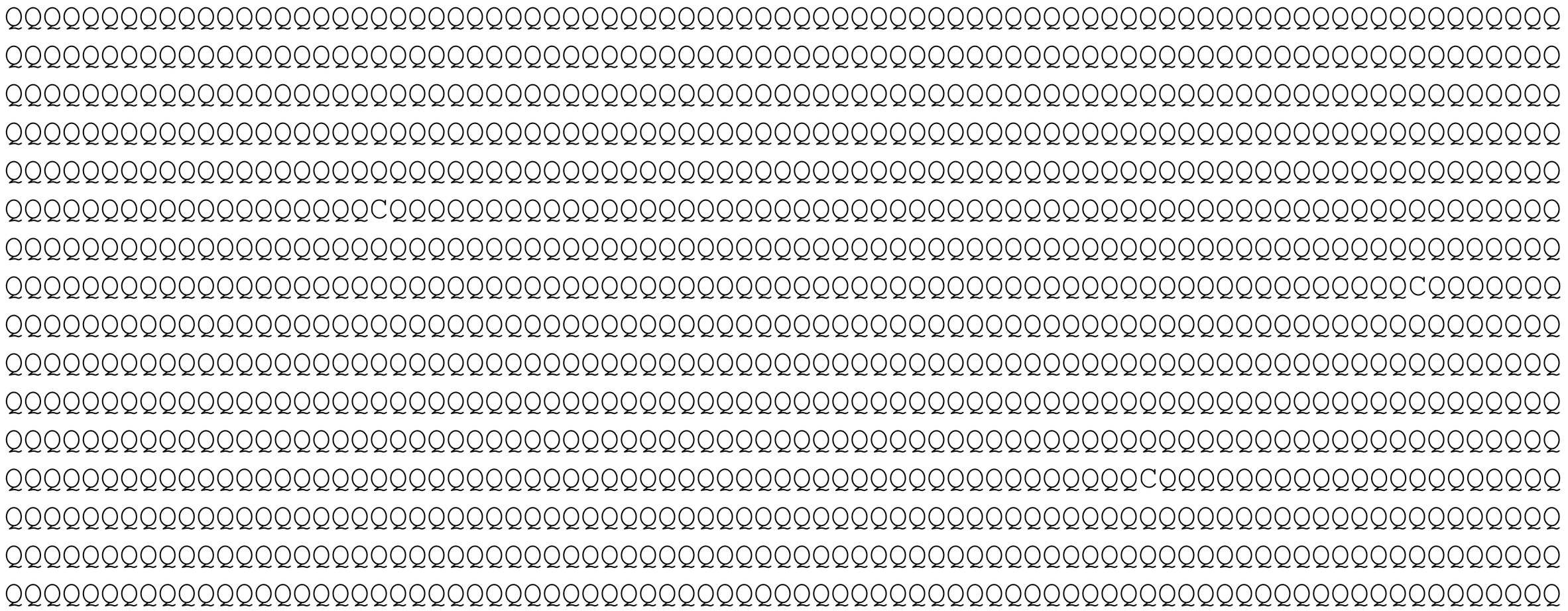
How many C's?

QCQQCQQQQCQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQCQQQQQQQQQQQQQQQQ
QQQQCQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQCQQQQQ

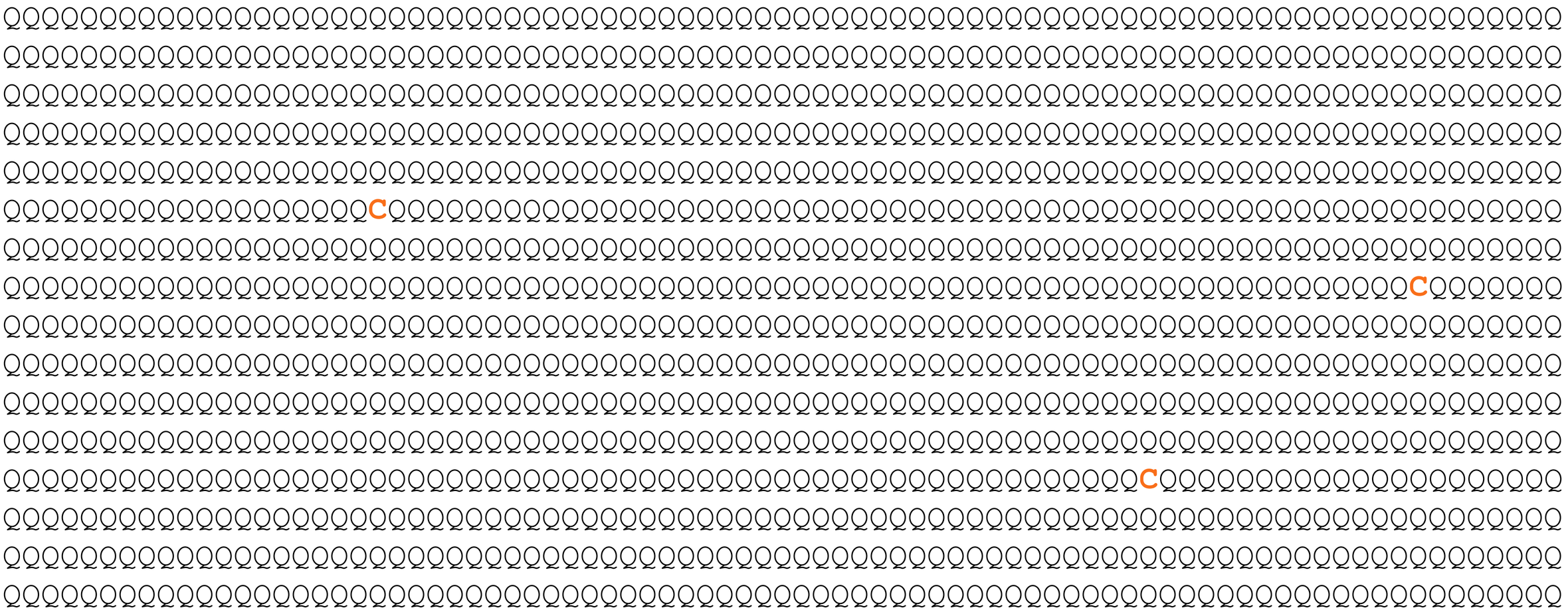
How many C's?


QCQQCQQQQCQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQCQQQQQQQQQQQQQQQQQ
QQQQCQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ
QQQQQQQQQQQQQQQQQQQQQQQQQQCQQQQQQ

How many C's?

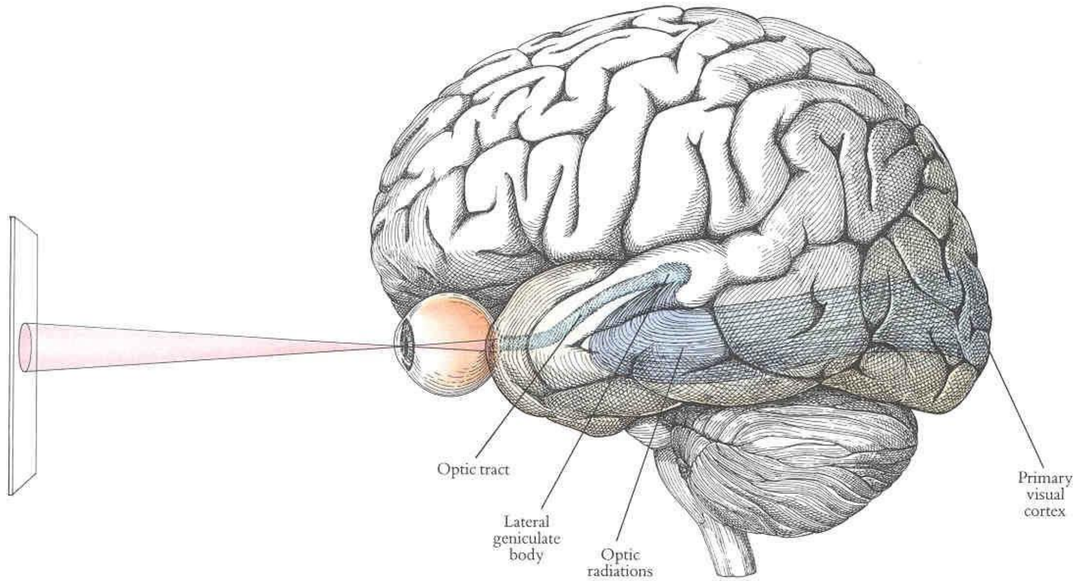


How many C's?



"… half of the human brain is devoted directly or indirectly to vision"

Mriganka Sur, MIT

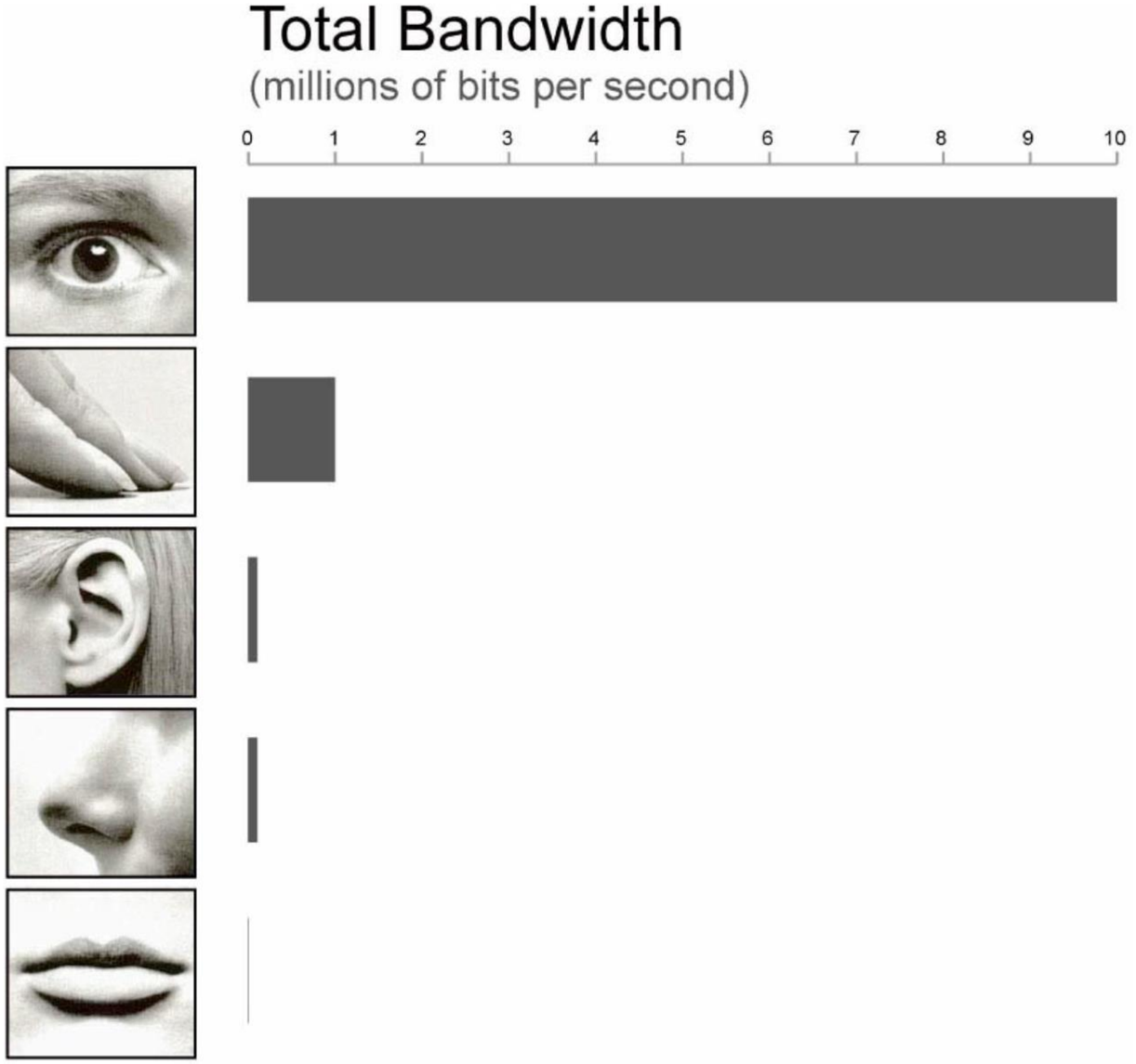


10 million bits per second

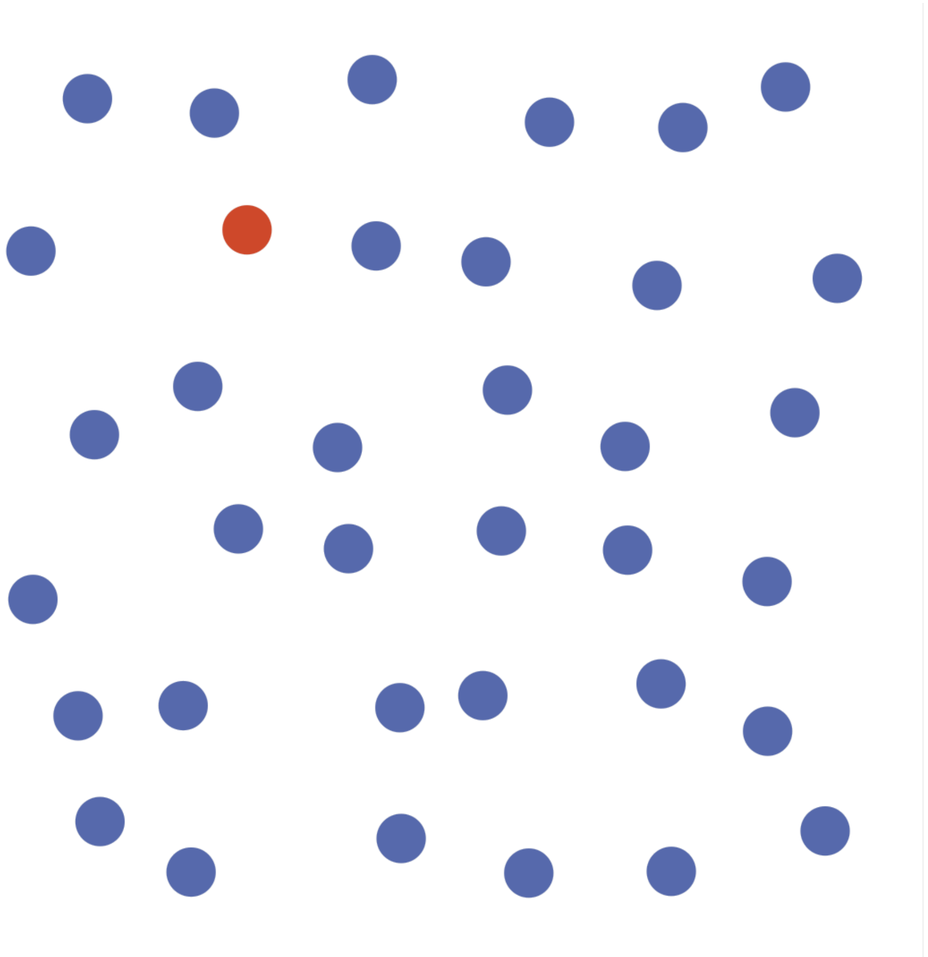
McLean & Freed, Current Biology (2006)

Image from "Approaches to the Mind: Introduction to Cognitive Science" Heather Bortfeld, Brown U.

It's efficient

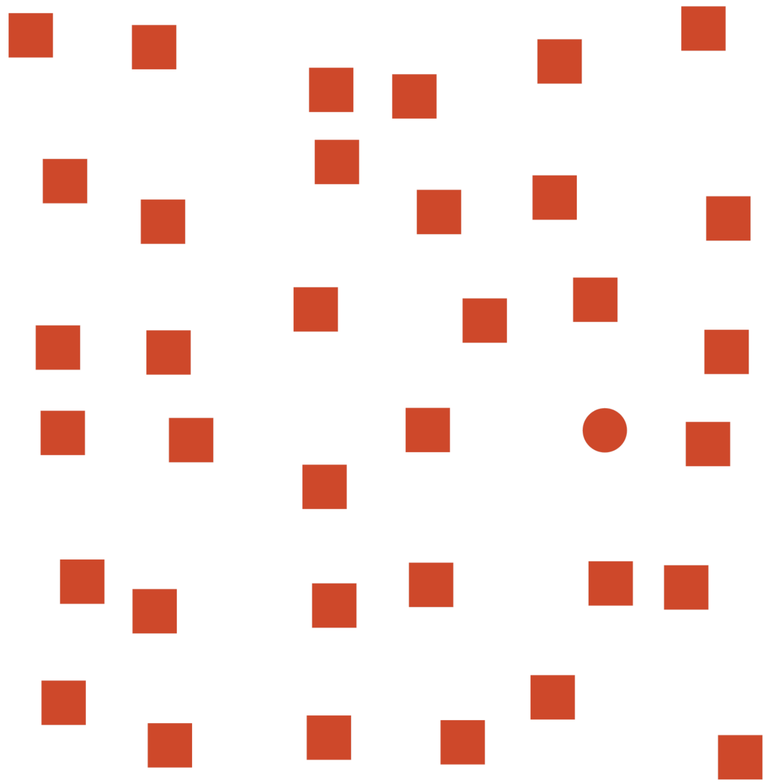


Pre-Attentive Processing



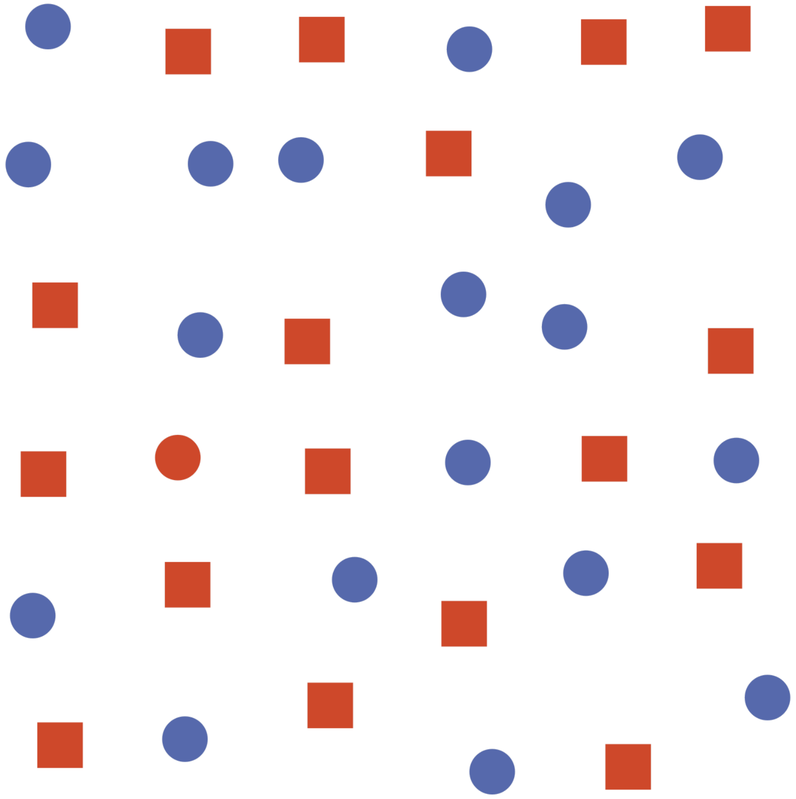
C. G. Healey and J. T. Enns,
“Attention and Visual Memory in Visualization and Computer Graphics,”
IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 7, pp. 1170–1188, 2012.

Pre-Attentive Processing



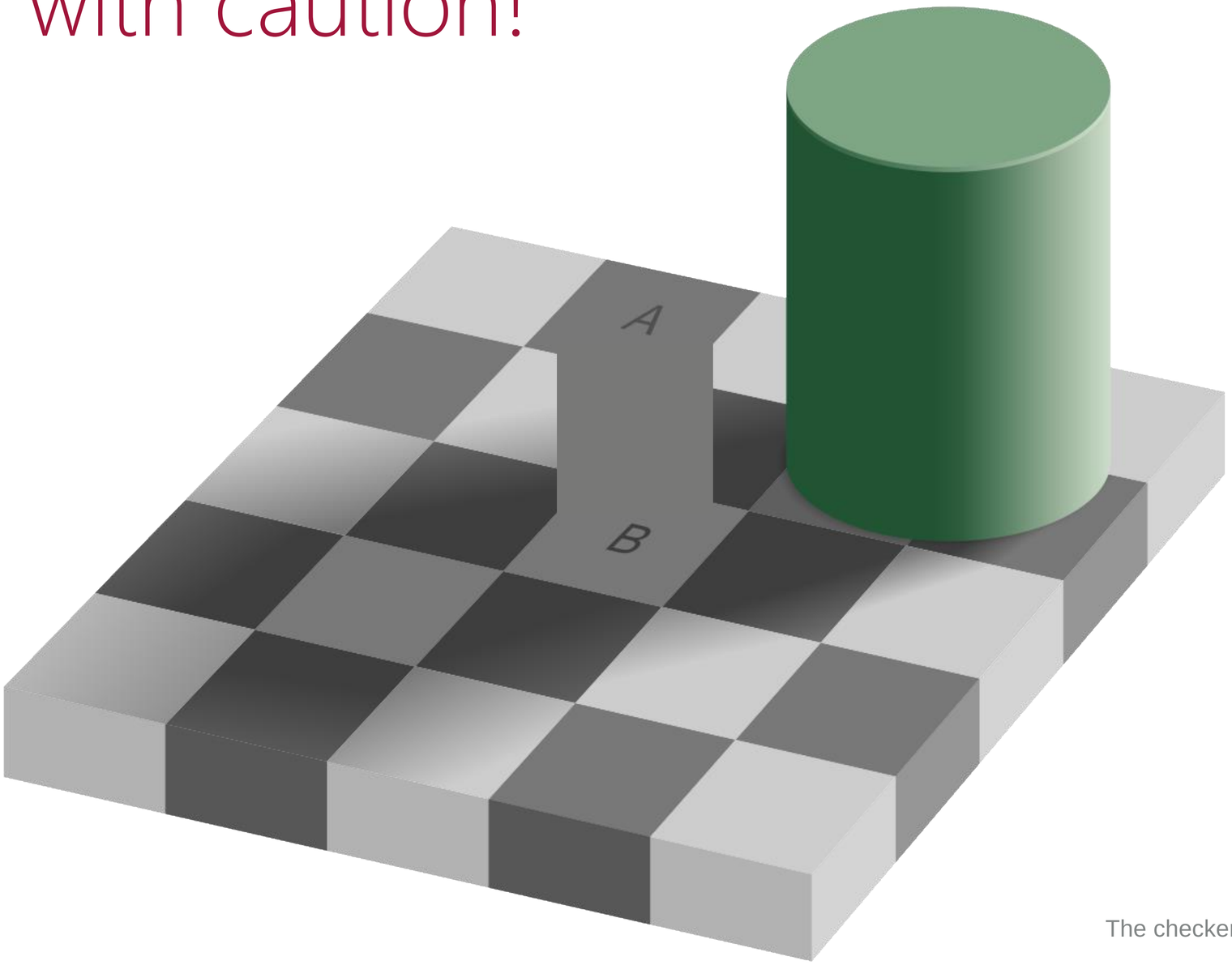
C. G. Healey and J. T. Enns,
“Attention and Visual Memory in Visualization and Computer Graphics,”
IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 7, pp. 1170–1188, 2012.

Pre-Attentive Processing



C. G. Healey and J. T. Enns,
“Attention and Visual Memory in Visualization and Computer Graphics,”
IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 7, pp. 1170–1188, 2012.

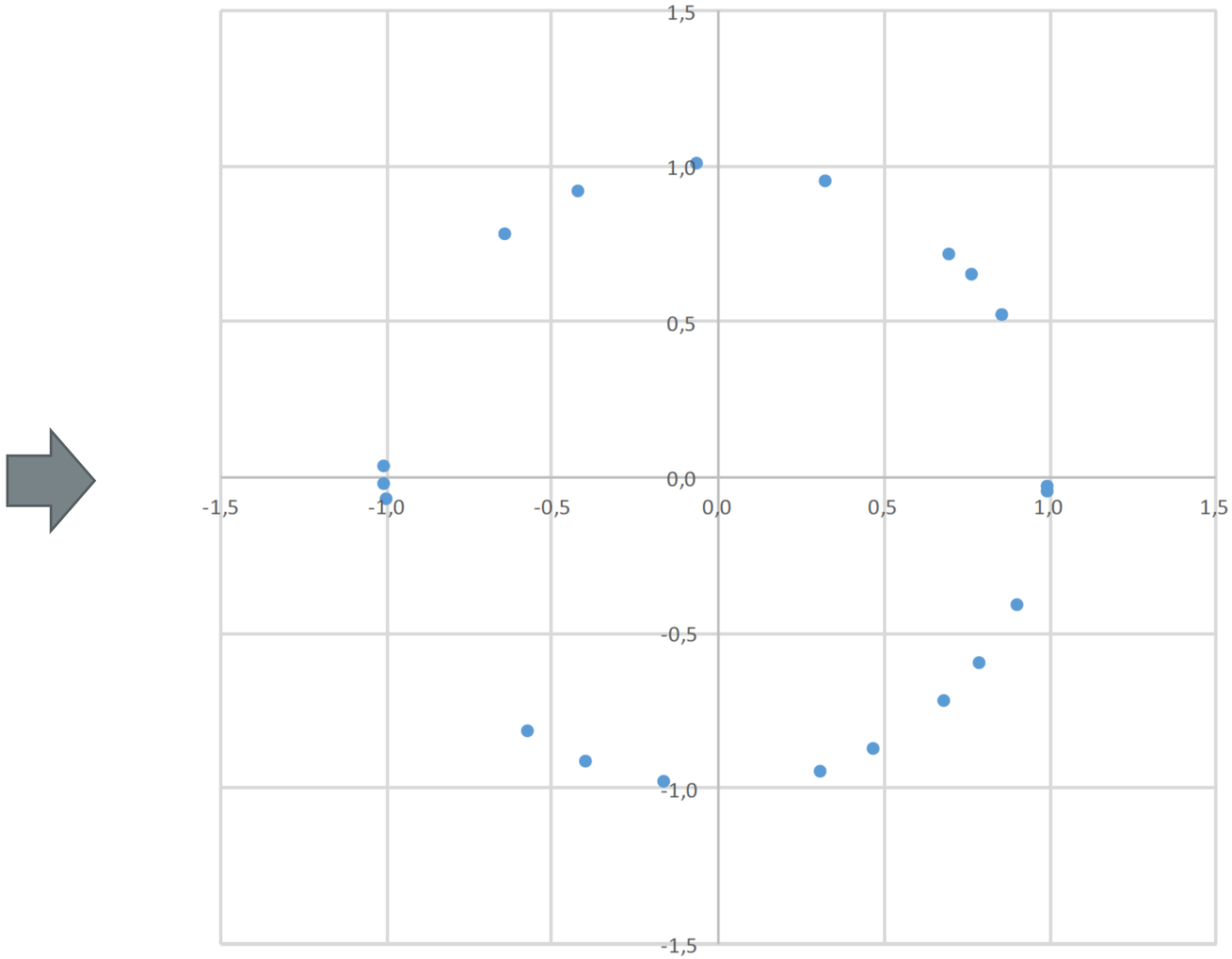
Use color with caution!



The checkerboard illusion of Edward Adelson

Spatialisation

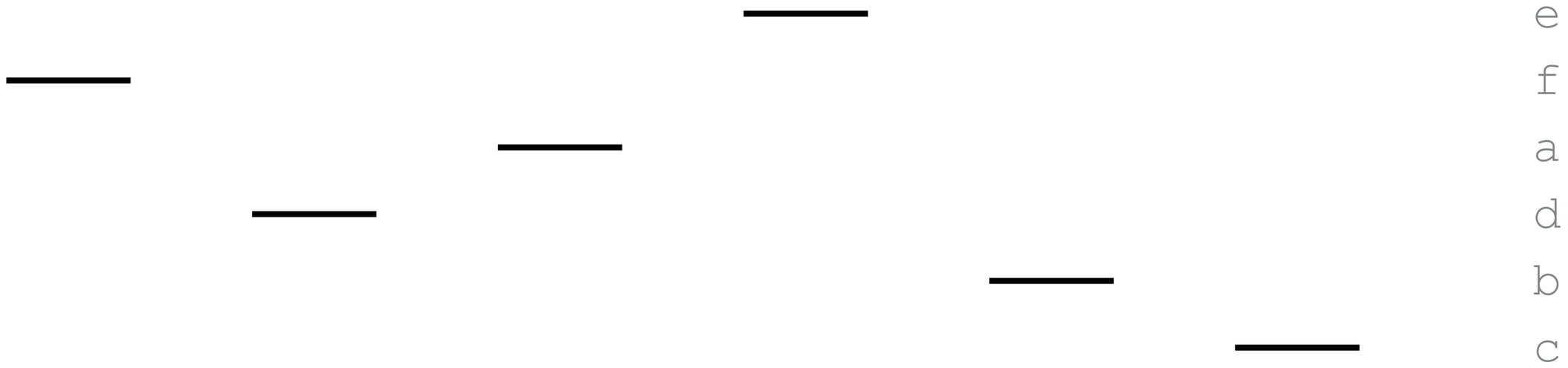
	X	Y
4,74481143	-1,0	0,0
2,82105932	0,3	-0,9
2,65059565	0,5	-0,9
5,59600326	-0,6	0,8
1,99963564	0,9	-0,4
1,0326167	0,9	0,5
3,5466094	-0,4	-0,9
4,63371265	-1,0	-0,1
3,29784409	-0,2	-1,0
2,22273521	0,8	-0,6
0,77896606	0,7	0,7
5,85675253	-0,4	0,9
3,744693	-0,6	-0,8
6,22601537	-0,1	1,0
1,62629688	1,0	-0,1
1,60465223	1,0	0,0
0,8747148	0,8	0,6
0,33548779	0,3	0,9
2,3868958	0,7	-0,7
4,68664757	-1,0	0,0



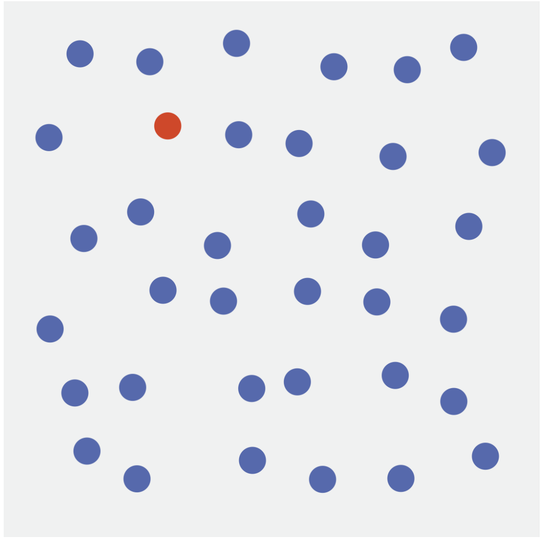
Visual Thinking

c < b f > d e > f b < d a > d a < f

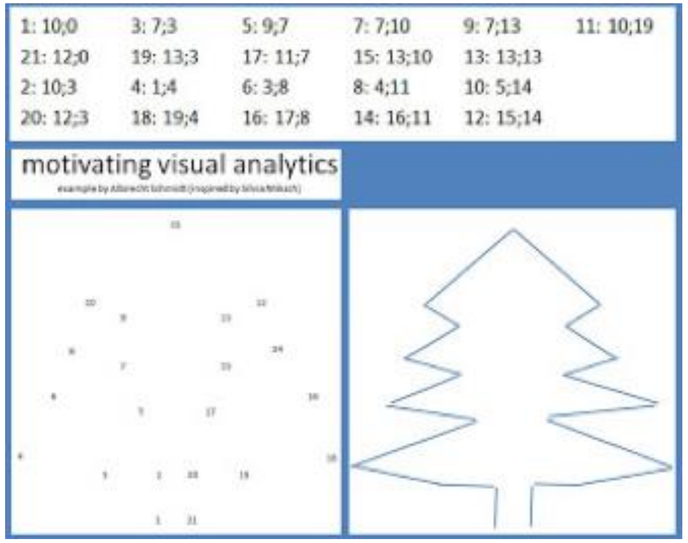
Visual Thinking



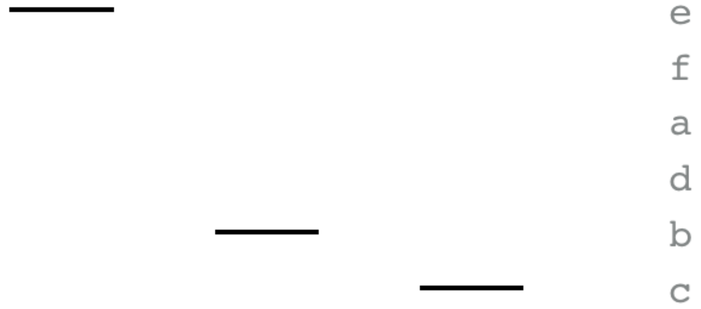
Summary: Why Use Visualization?



Visual Search



Recognition



Perceptual Inference

When is Visualisation Required?

When is Visualisation Required?

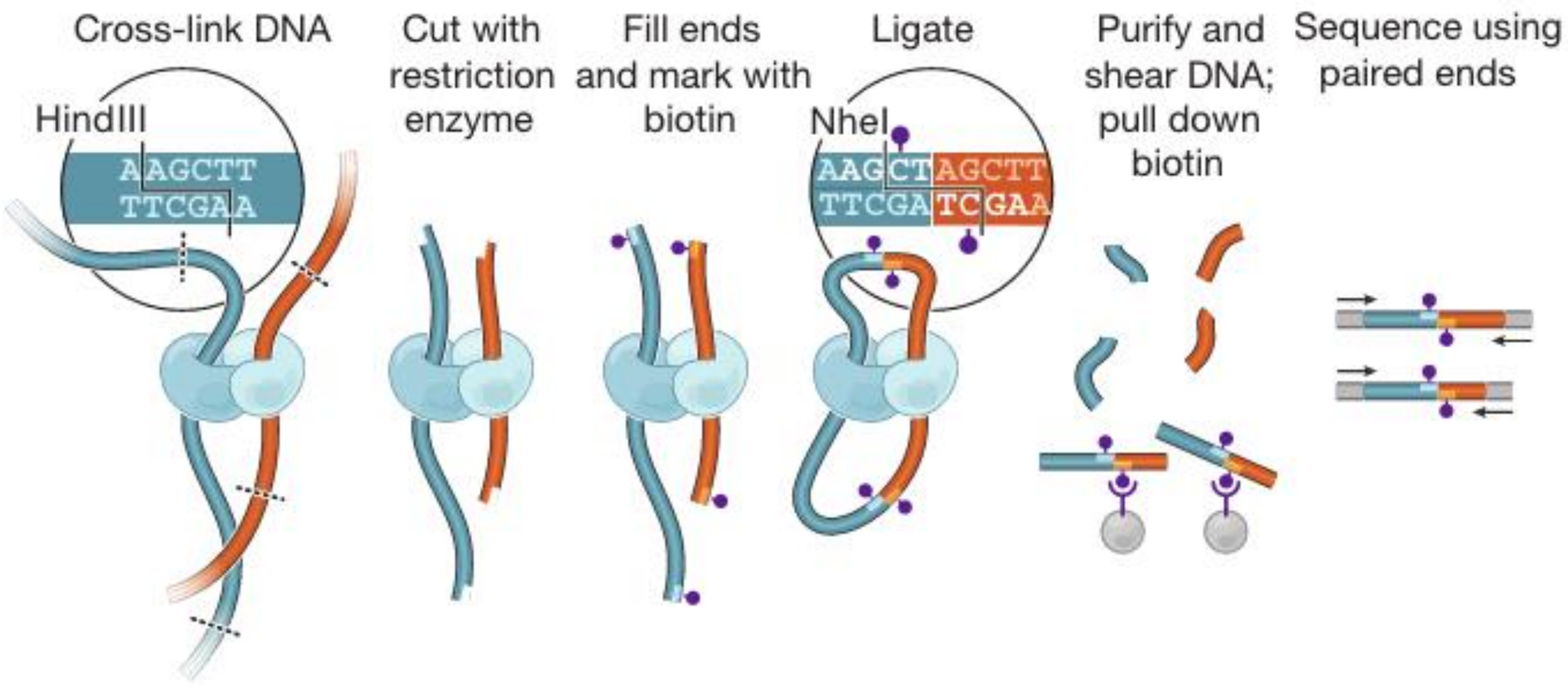
- Problem is not sufficiently well defined
- Question unclear or too many possible questions
- No automatic algorithm available

→ **Needs a human in the loop!**

To explore the data

S. I. O'Donoghue et al., "Aquaria: simplifying discovery and insight from protein structures," Nat Meth, vol. 12, no. 2, pp. 98–99, 2015.

To explain and communicate



Lieberman-Aiden, E., et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome, 2009

Memory Loss

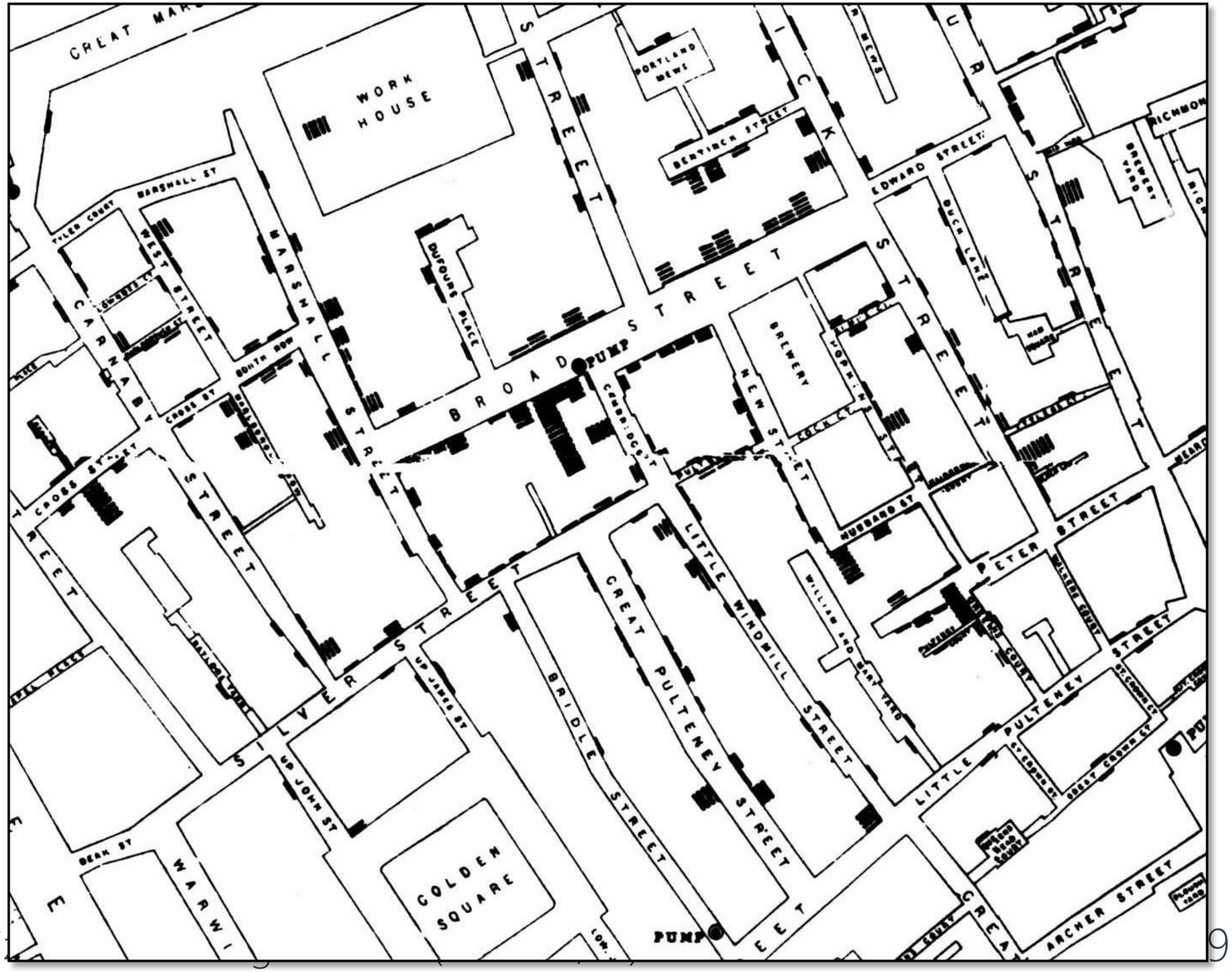
Difficulty Thinking

Confusion

Unusual Behaviour

To analyse and discover

1854
John Snow
The map of cholera



To test and form hypotheses

Anscombe's Quartet: Raw Data

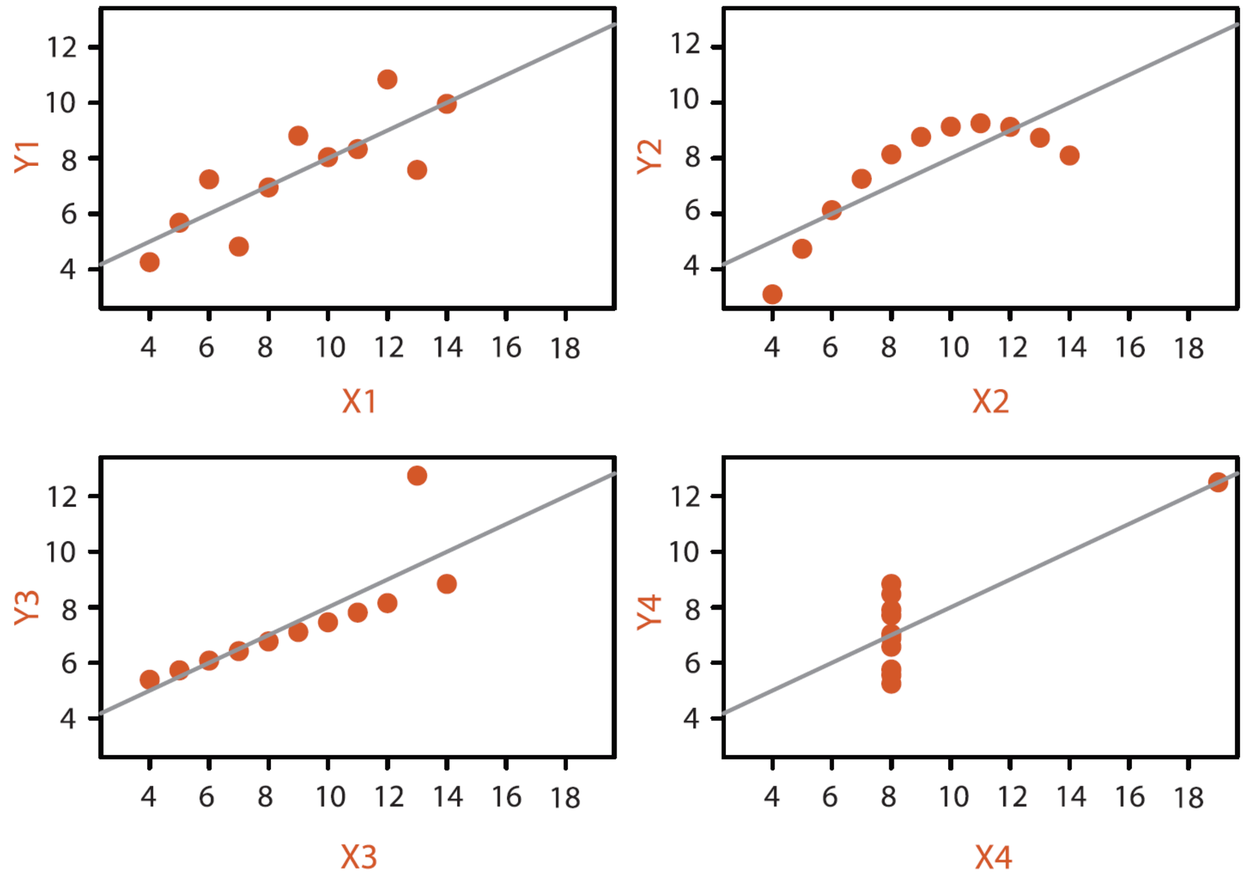
	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

Anscombe, F. J. Graphs in statistical analysis., 1973

To test and form hypotheses

Oversimplification →

Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	



Anscombe, F. J. Graphs in statistical analysis., 1973 Munzner, 2014

To monitor automated systems

- Multi-monitor setup (more screen space/resolution)

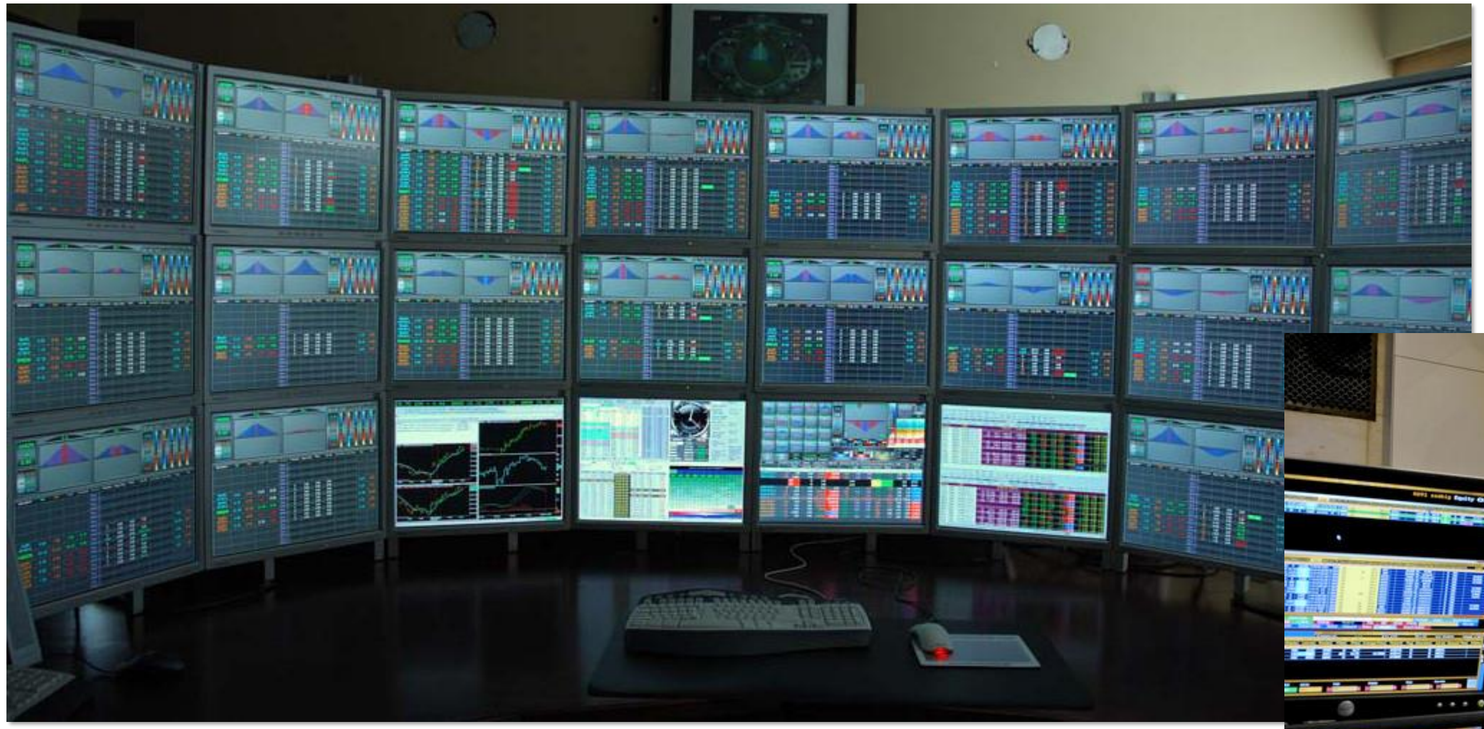
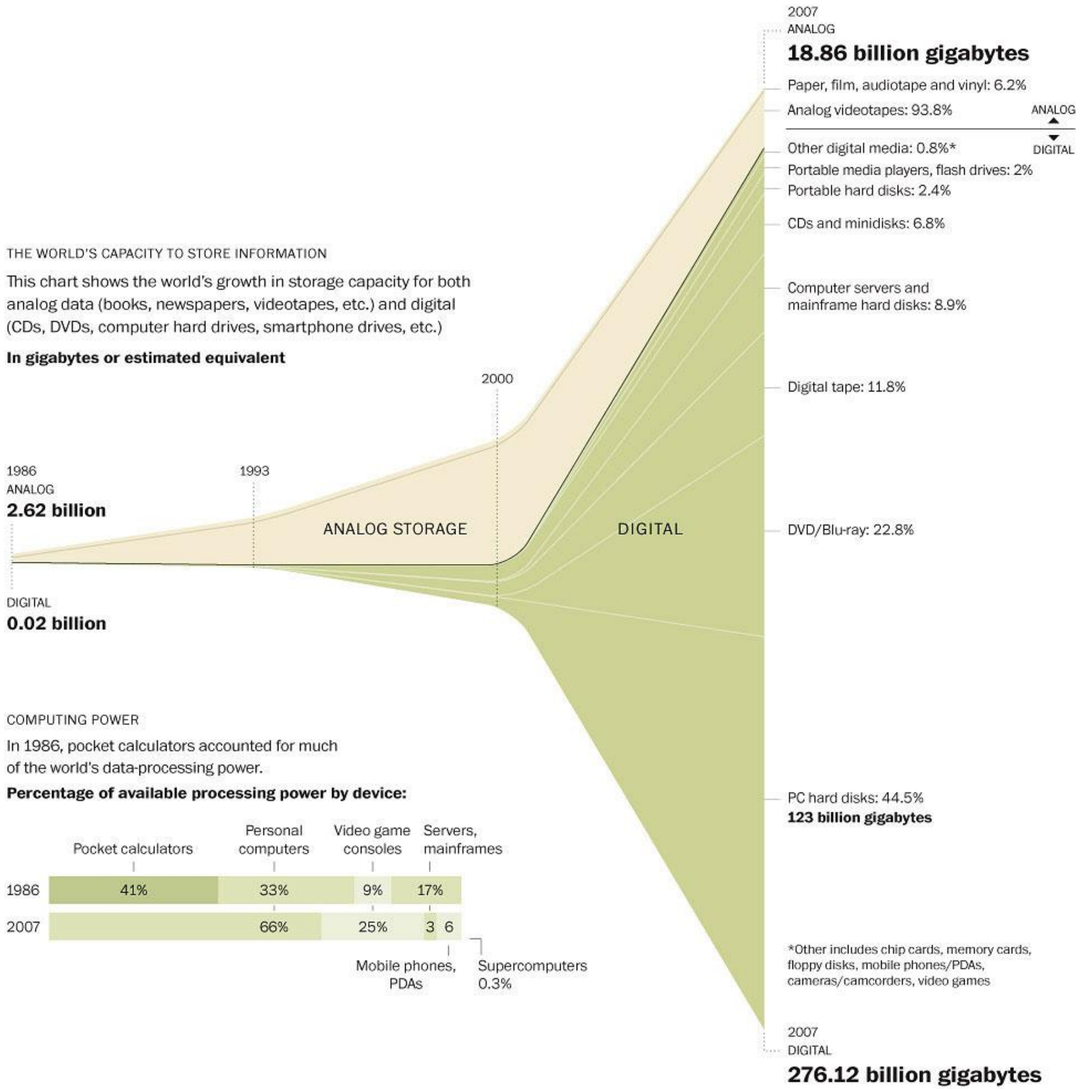
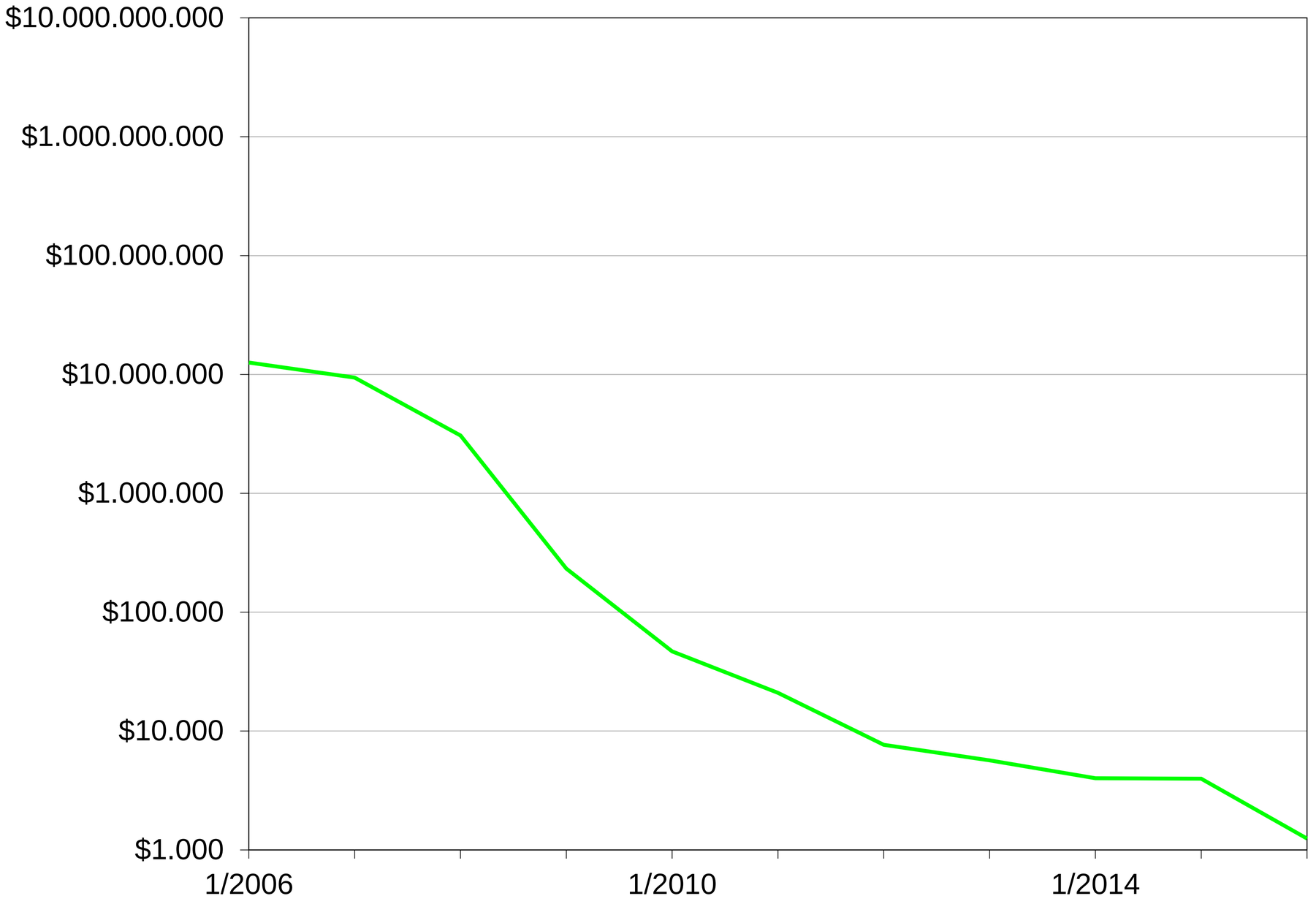


Image sources:
https://makemoneyinlife.com/
https://multimonitorcomputer.com/
https://kennethfriedman.org/

Why do we need a computer?

Why do we need a computer?



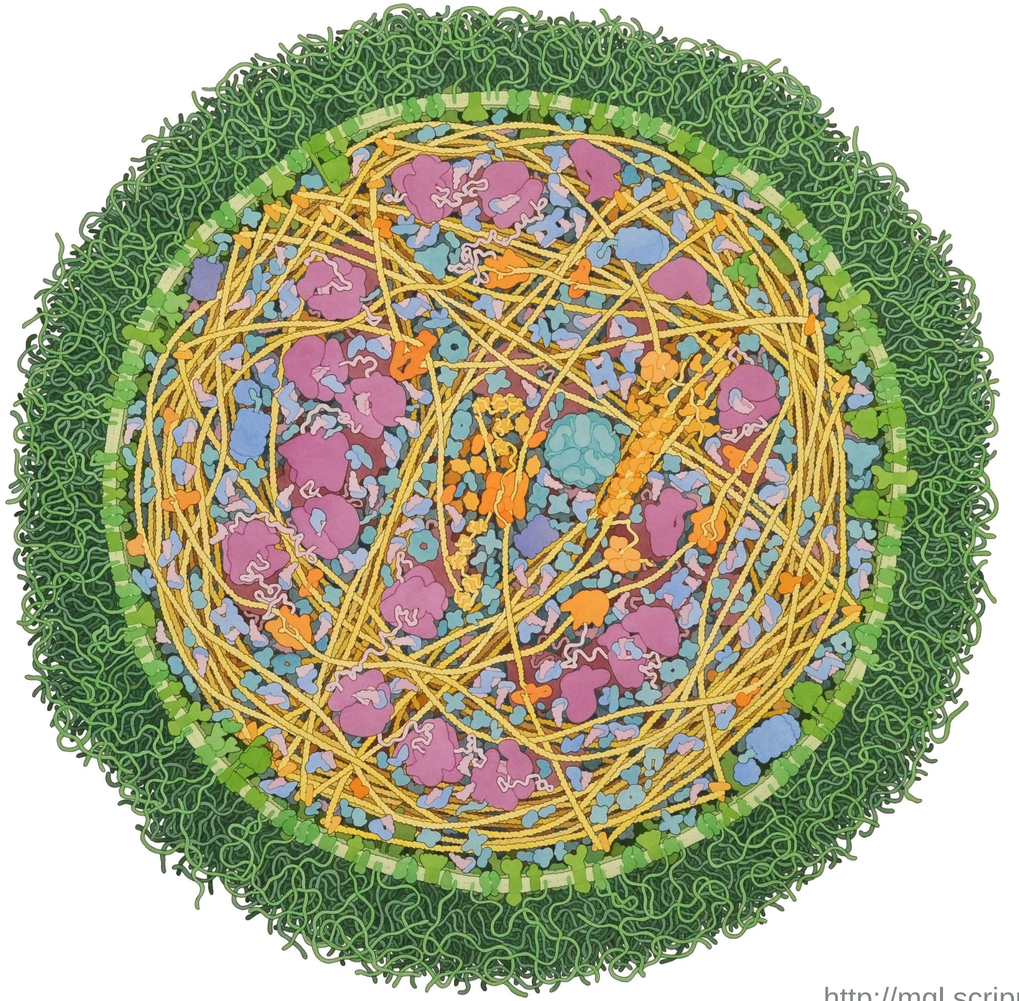


Cost per Genome

Why do we need a computer?

- Large datasets
- Automatic generation
- Interactivity

Hand-drawn or made using computer?



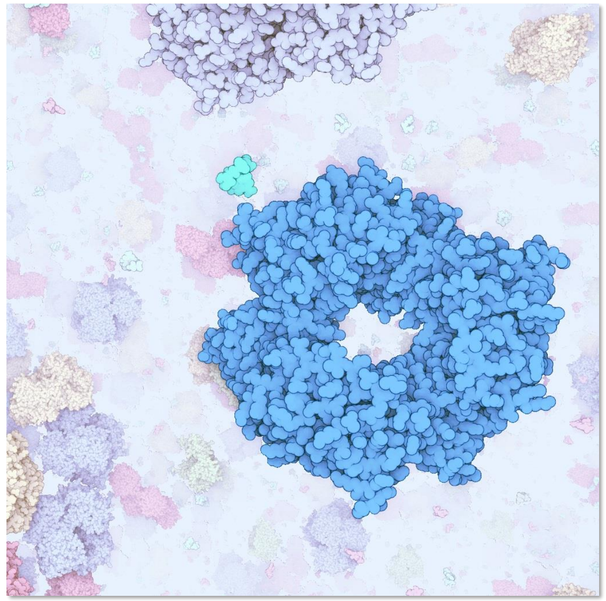
http://mgl.scripps.edu/

David S. Goodsell

Video

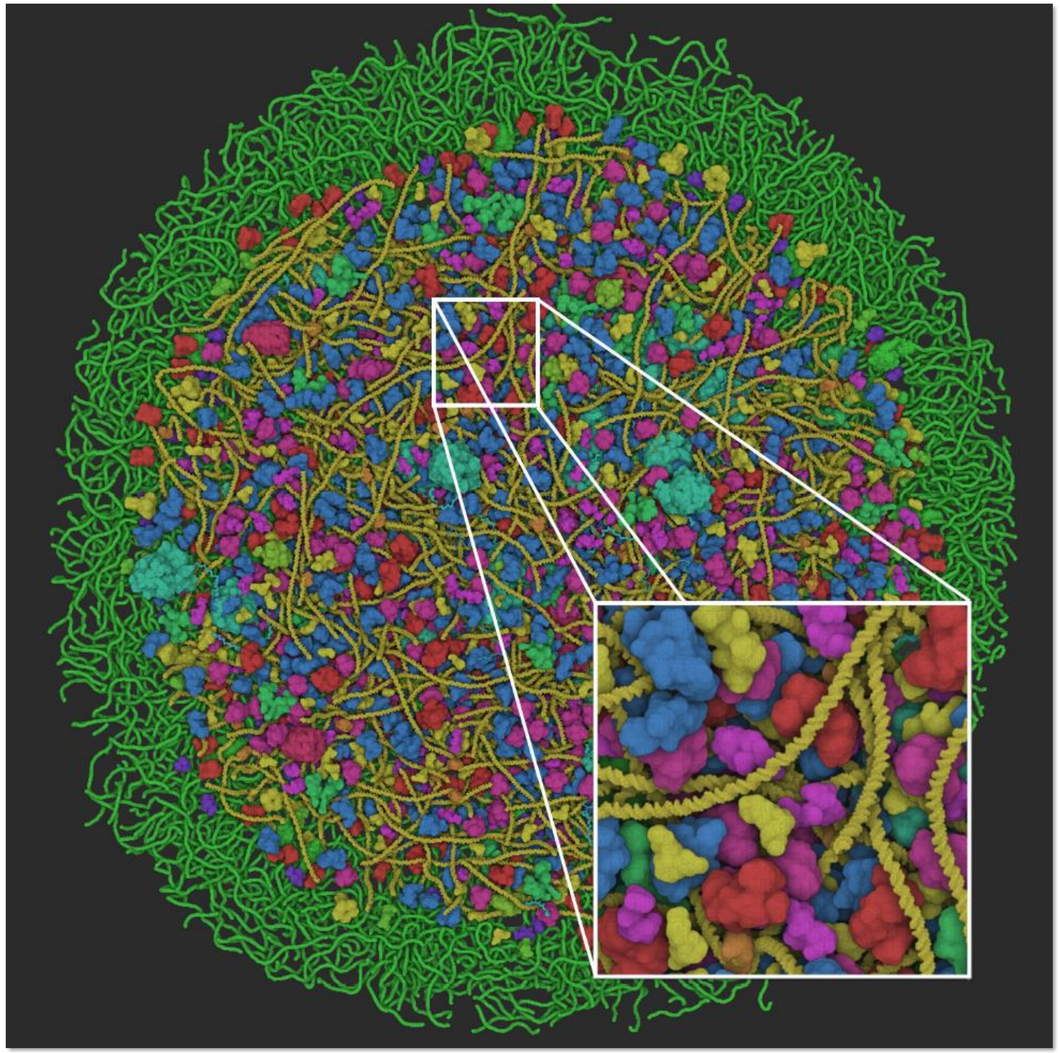
(link in extra material)

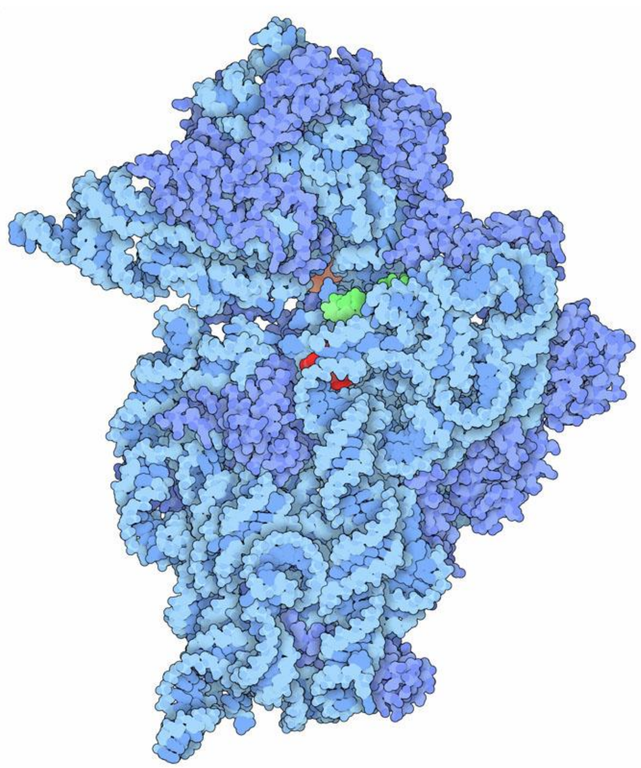
Hand-drawn or made using computer?



▲ M. Le Muzic, J. Parulek, A. Stavrum, I. Viola: *Illustrative Visualization of Molecular Reactions using Omniscient Intelligence and Passive Agents*.
Computer Graphics Forum 33(3):141-150, 2014.

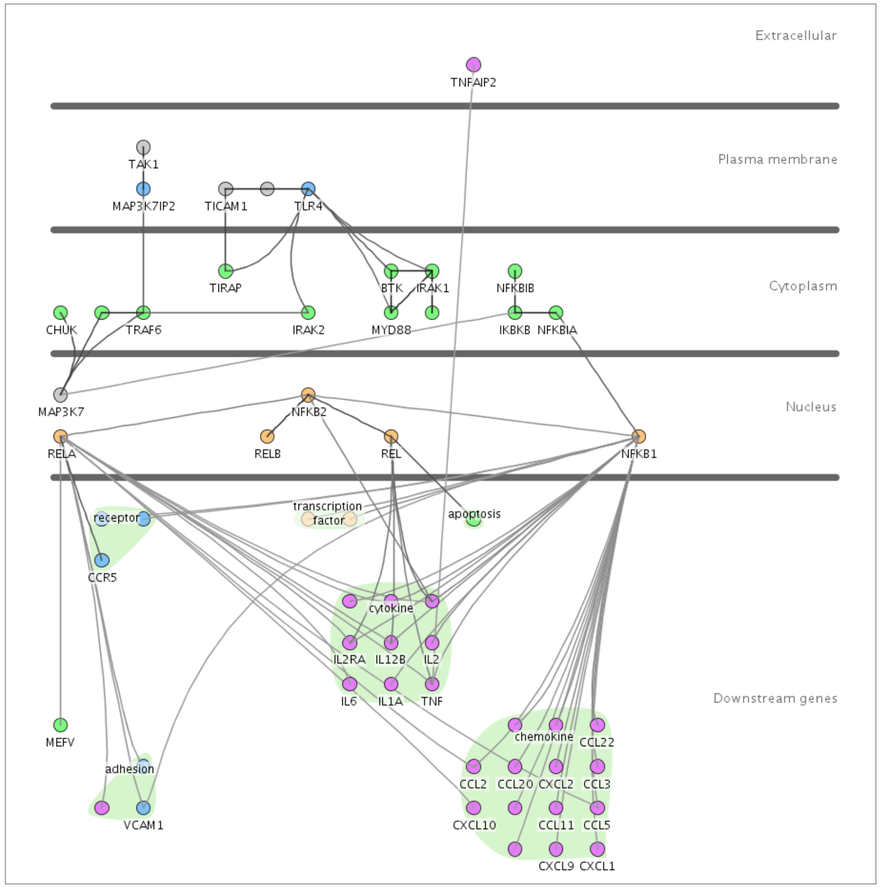
► M. Le Muzic, L. Autin, J. Parulek, I. Viola: *cellVIEW: a Tool for Illustrative and Multi-Scale Rendering of Large Biomolecular Datasets*.
EG Visual Computing for Biology and Medicine, 2015.





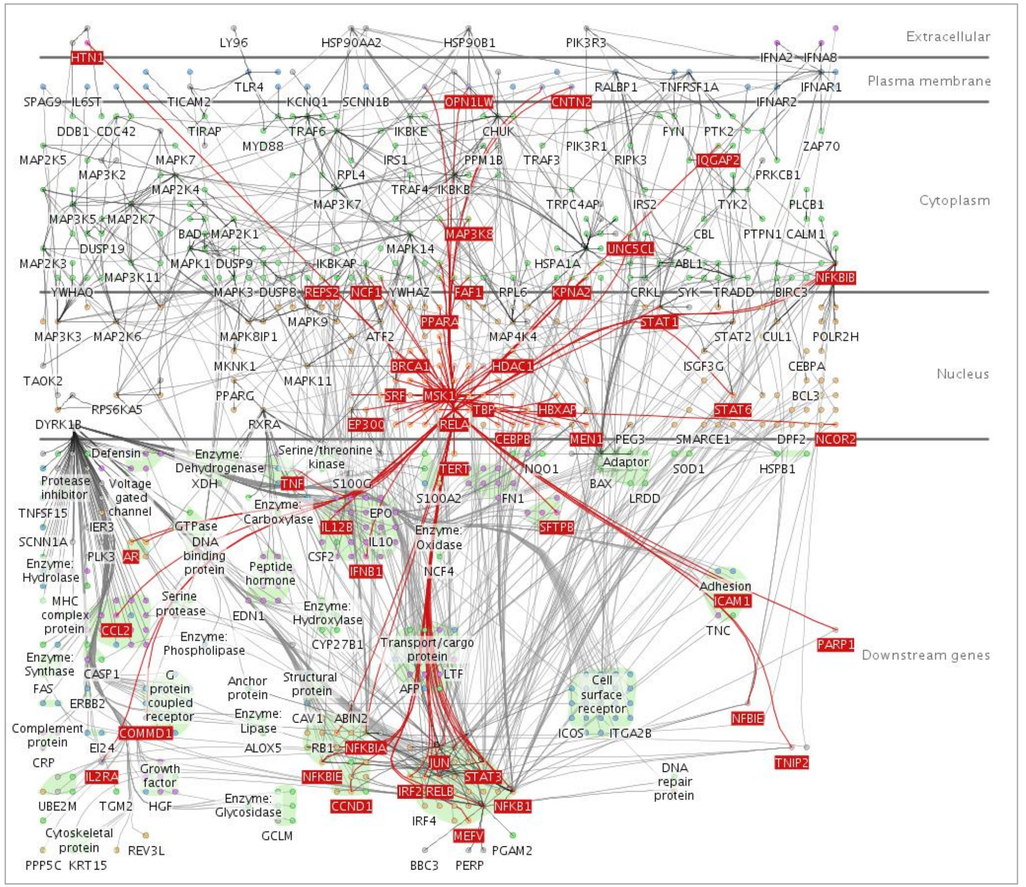
▲ D. S. Goodsell, S. Dutta, C. Zardecki, M. Voigt, H. Berman, S. Burley: *The RCSB PDB "Molecule of the Month": Inspiring a Molecular View of Biology*.
PLoS Biol 13: e1002140, 2015.

Hand drawn



57 nodes, 74 edges

Automatically generated



760 nodes, 1269 edges

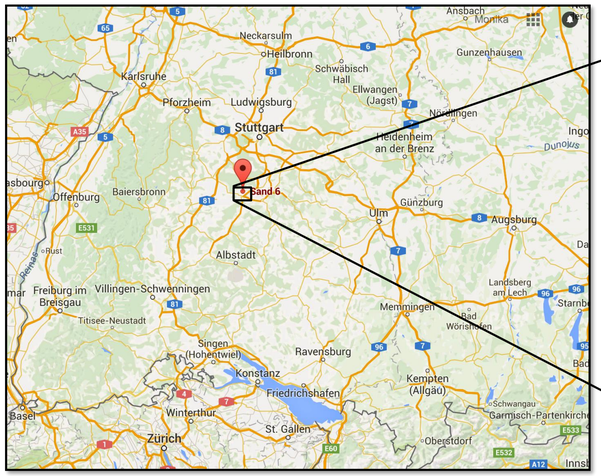
Barsky et al.: Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. Bioinformatics, 2007.

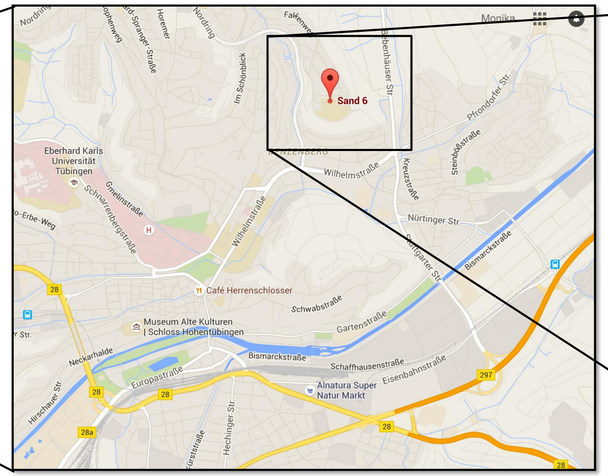
Why do we need a computer?

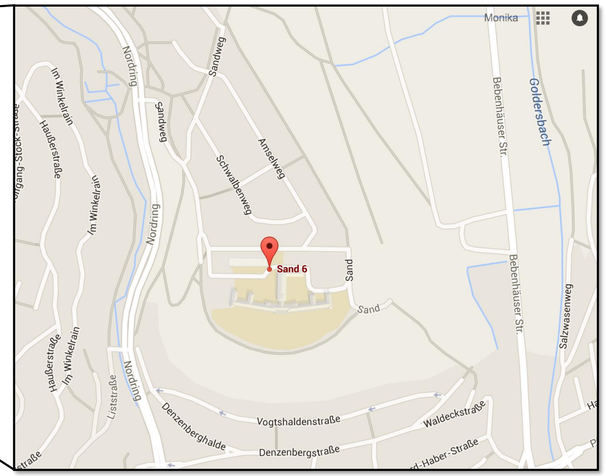
- Large datasets
- Automatic generation
- Interactivity

When make it interactive?

- Complex data
- Multiple levels of detail
- Multiple aspects
- Multiple representations







Ben Shneiderman's Mantra

Overview first,

Zoom and Filter,

Details on demand.

B. Shneiderman,

"The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations,"

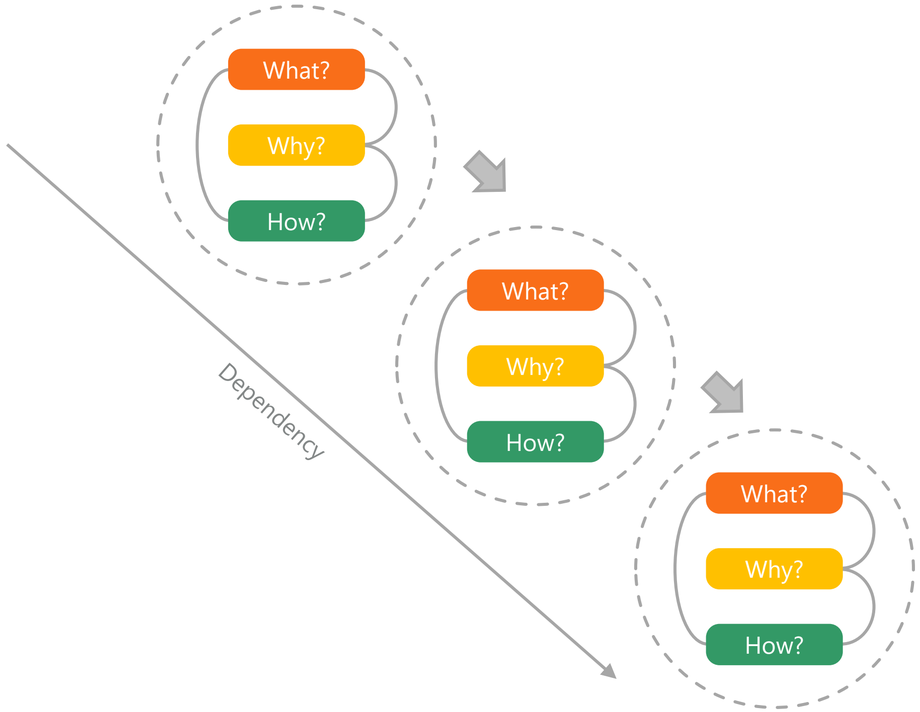
in Proceedings of the IEEE Symposium on Visual Languages, 1996.

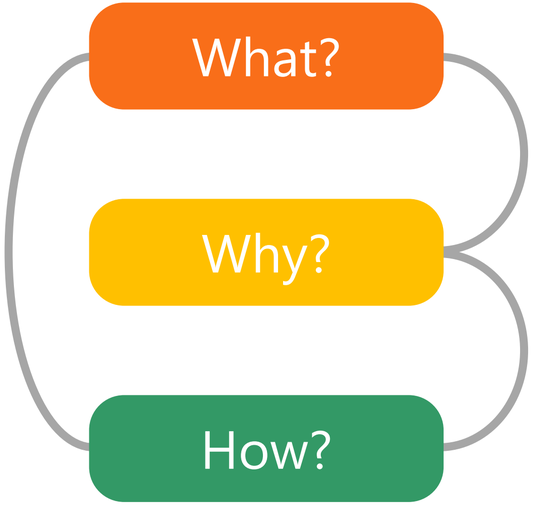
Design Considerations

- Different tasks might require different visualizations
- Alternative solutions
- Effectiveness over beauty
- Validation is crucial, but difficult
- Look into existing visualizations (tools)

Visualisation Analysis and Design

- What data user sees
- Why the user wants to use the tool
- How is it constructed in terms of the design choices





What are the limitations?

1. Computational Capacity

- Older systems, newer (larger) data
- Data size exceeding computational memory
- Interactivity requires fast computations
- **Scalability**

2. Human perceptual and cognitive capacity

Change Blindness

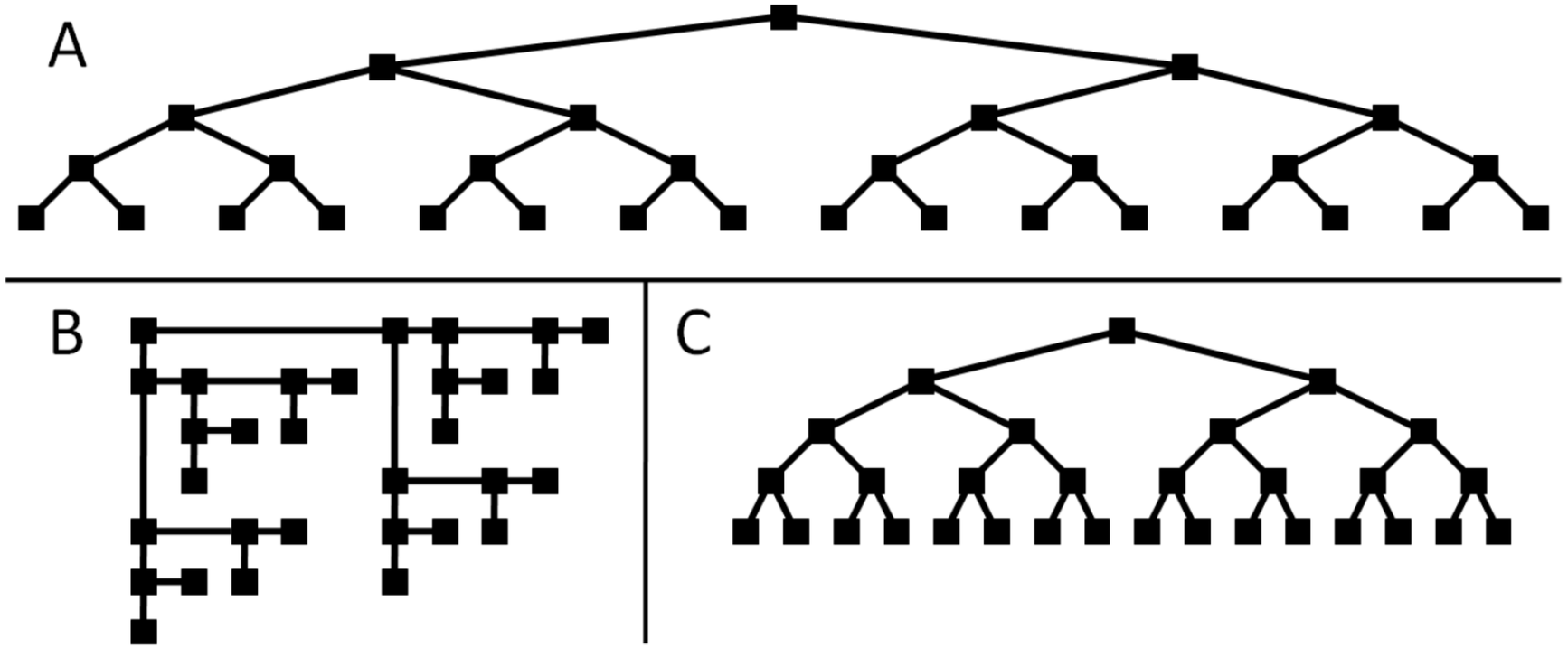
Human perceptual and cognitive capacity

- Change Blindness

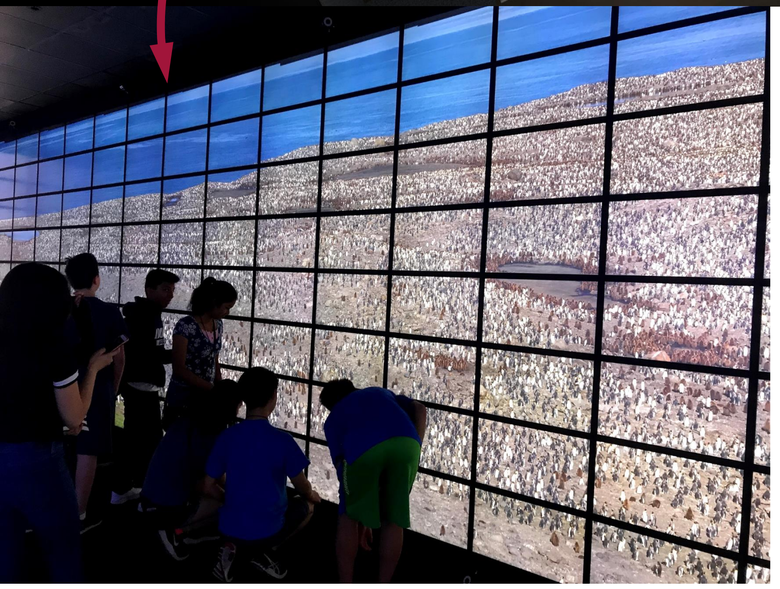
Human perceptual and cognitive capacity

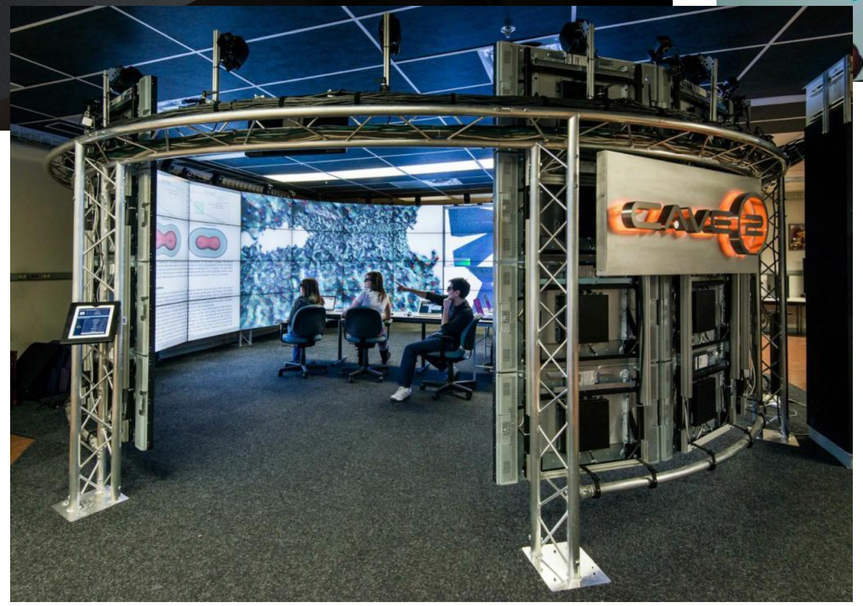
- Change Blindness
- Attention

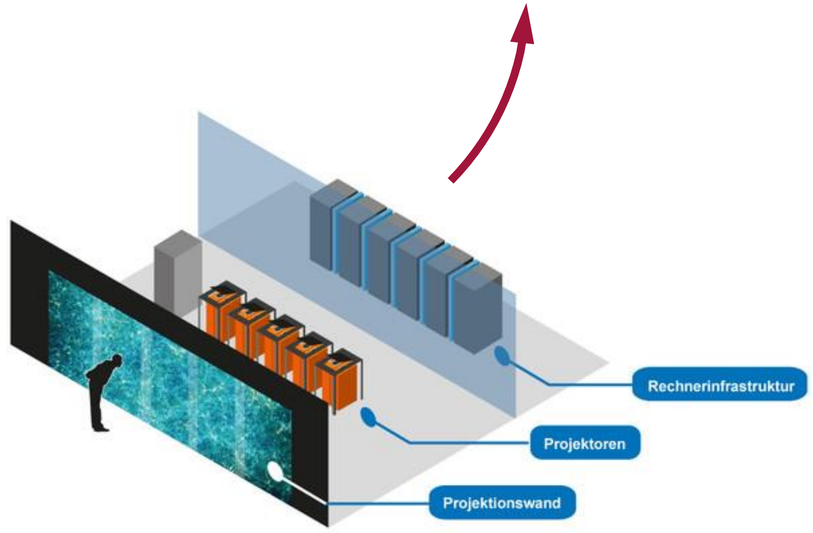
3. Display Capacity



McGuffin, M. J., and Jean-Marc R. "Quantifying the space-efficiency of 2D graphical representations of trees., 2010







▲ CAVE 2, University of Illinois Chicago

◄ Reality Deck, Stony Brook University NY

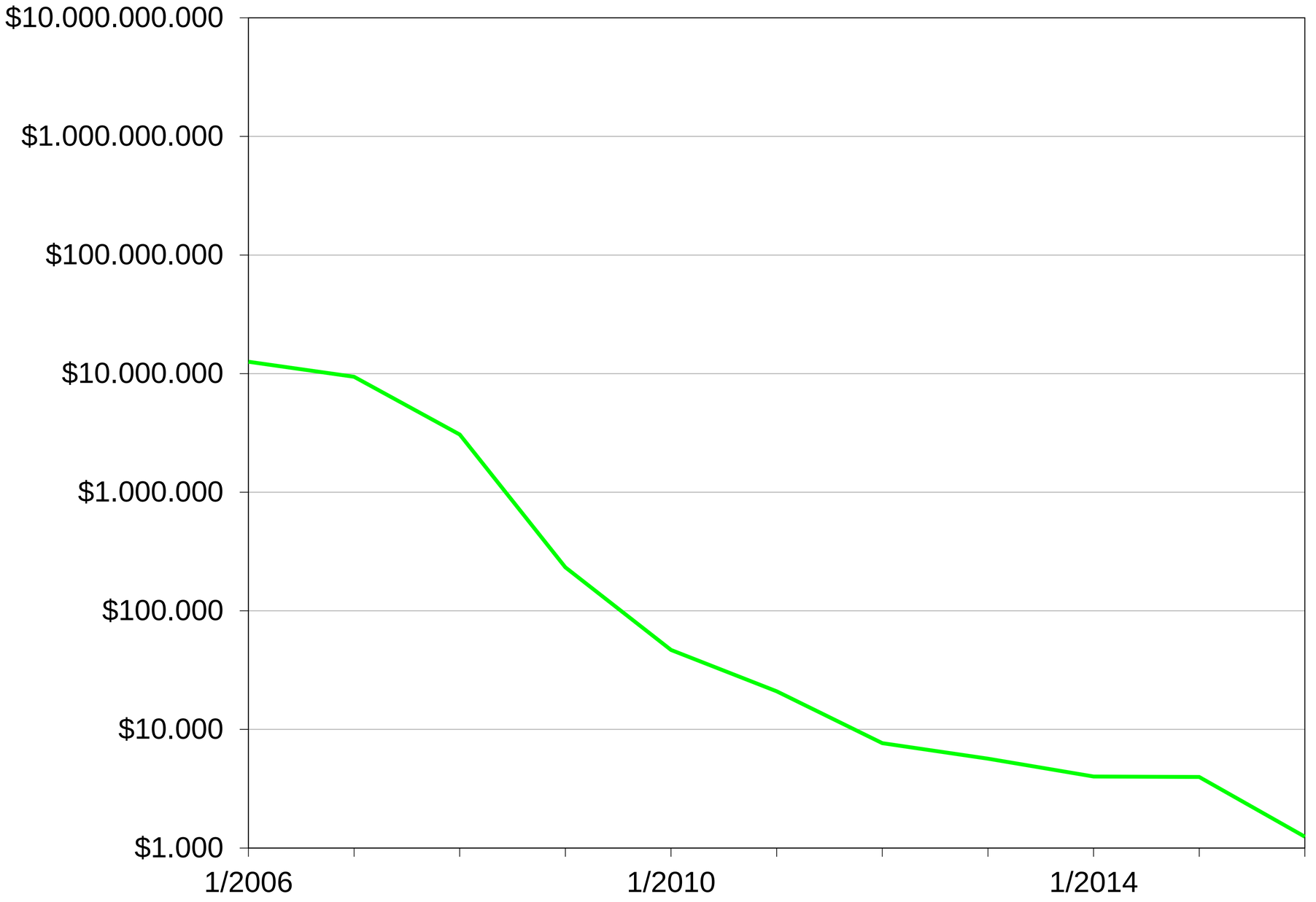
▲ VVand, Universität Stuttgart

Limitations

- Computational capacity
- Human perceptual and cognitive capacity
- Display capacity

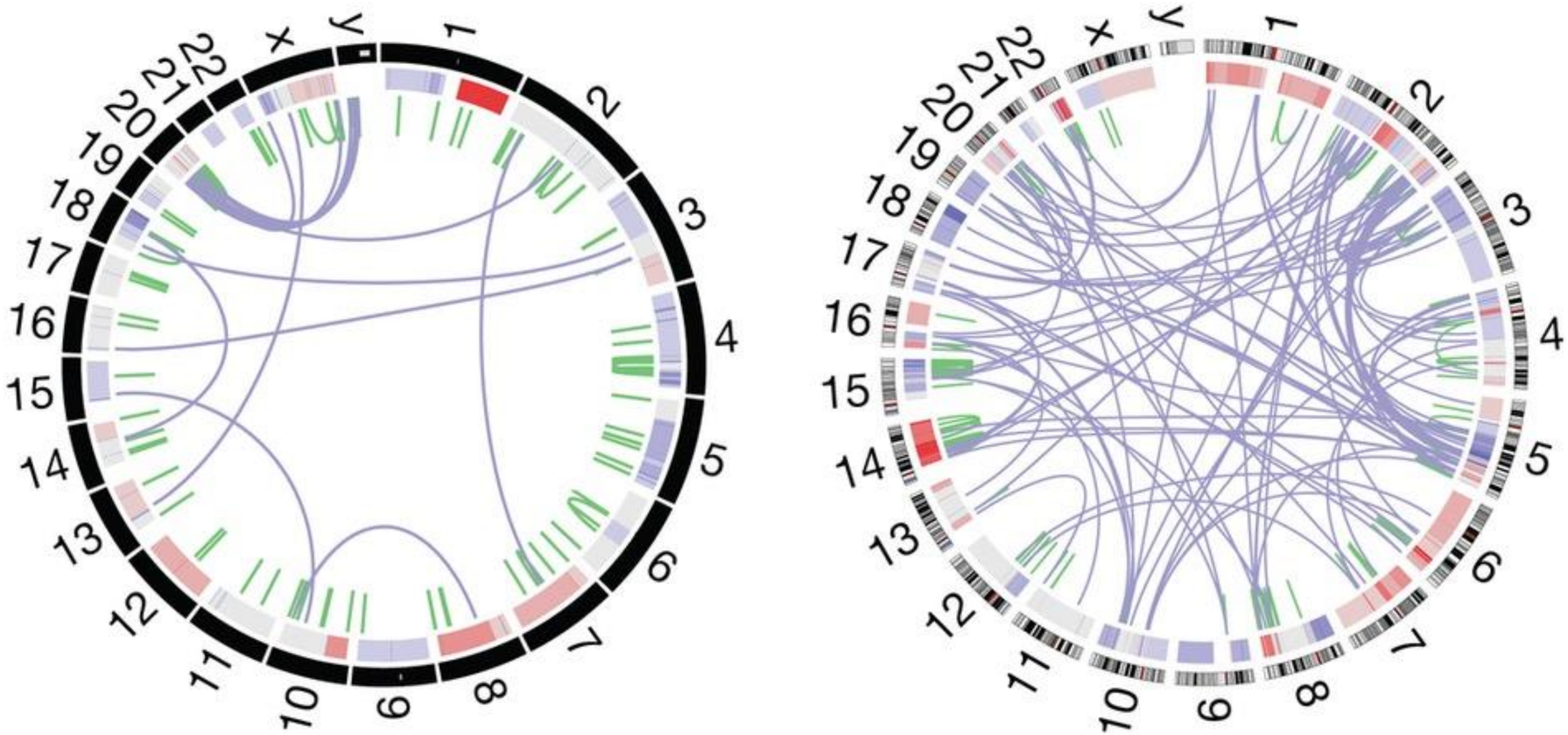
INTRODUCTION

Visualization in Biology



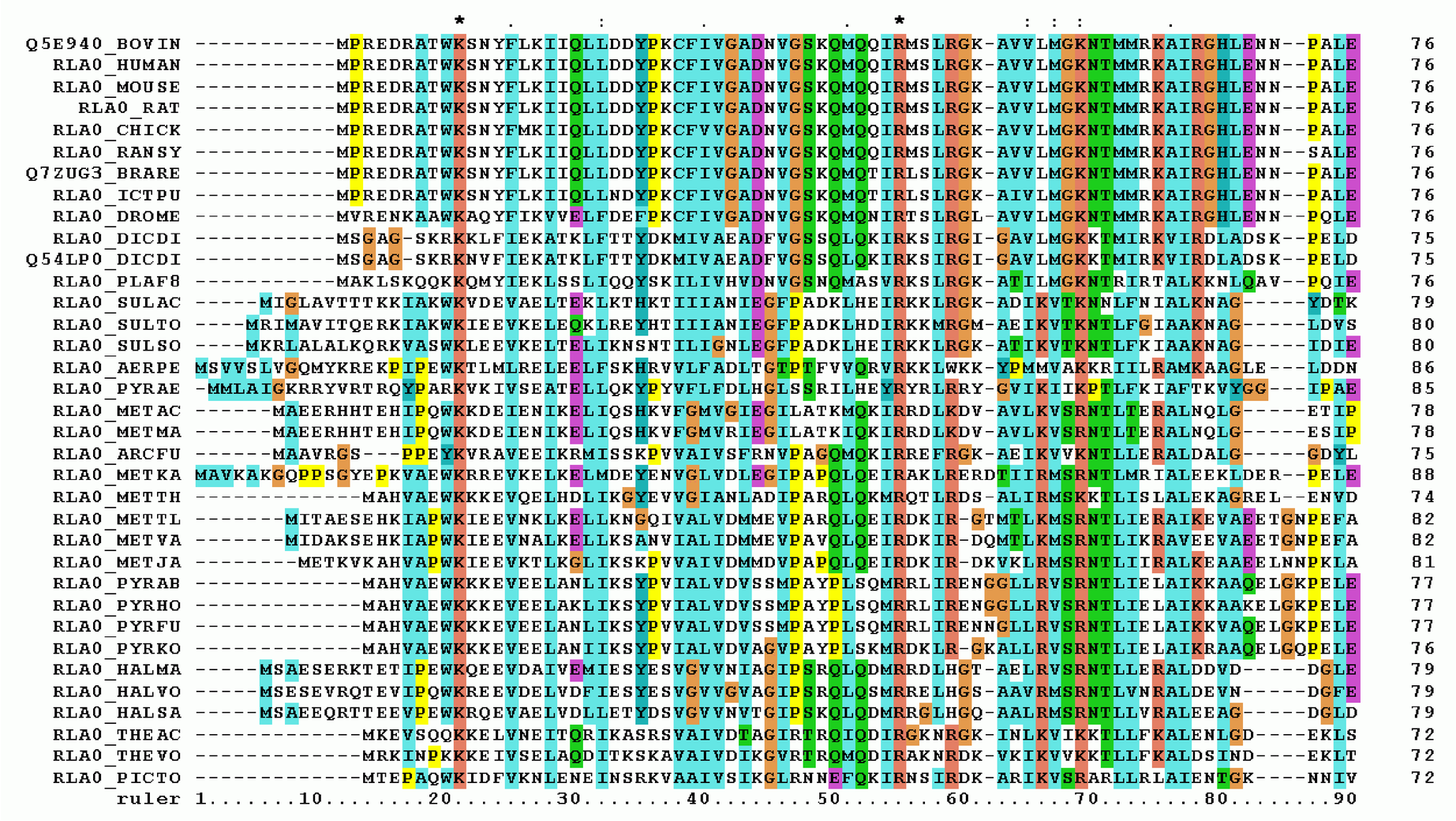
Cost per Genome

Genomes



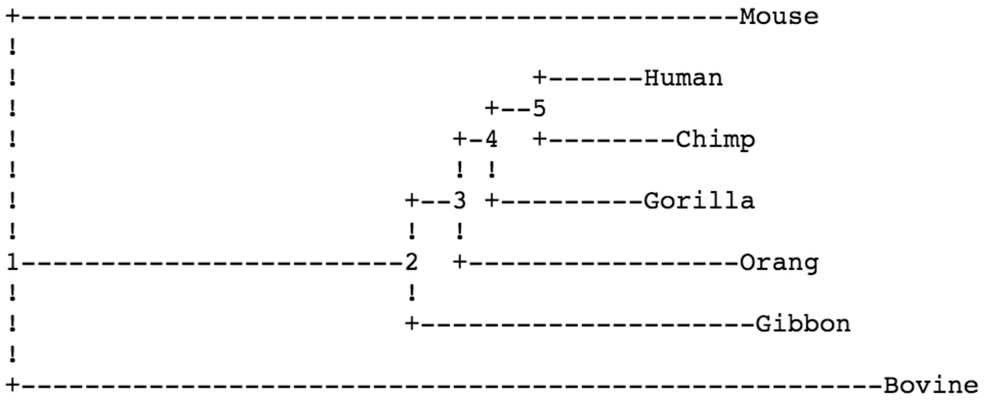
Imielinski, M., et al. "Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing., 2012

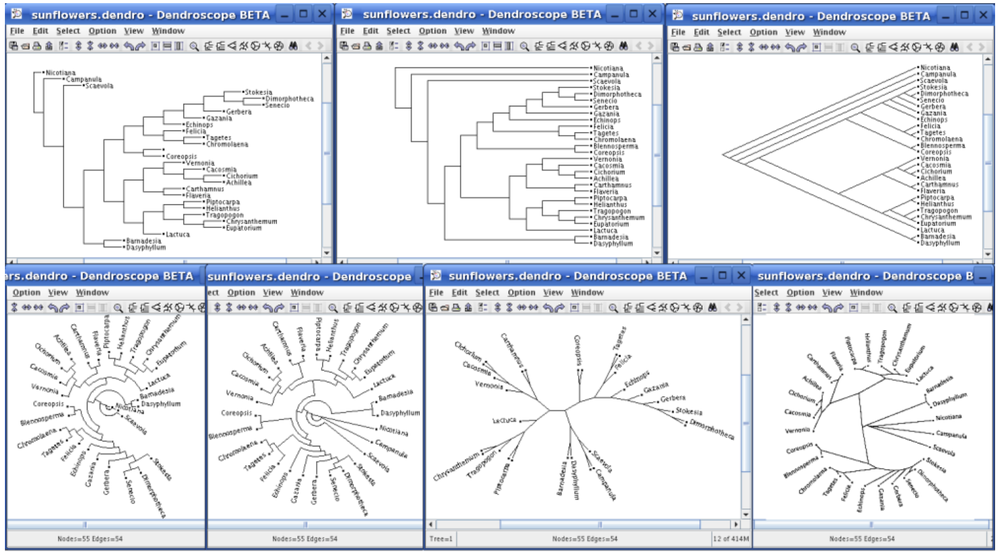
Multiple Sequence Alignments



en.wikipedia.org

Phylogenies

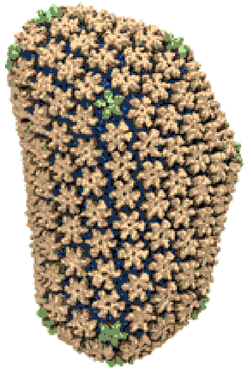


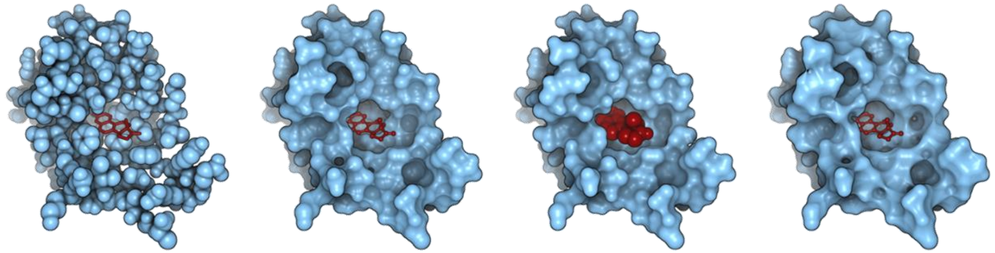


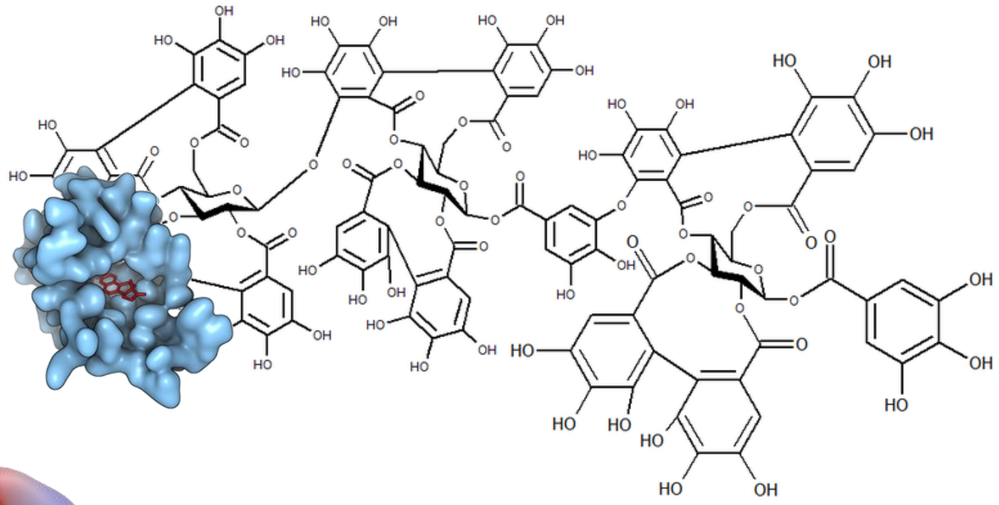
evolution.genetics.washington.edu

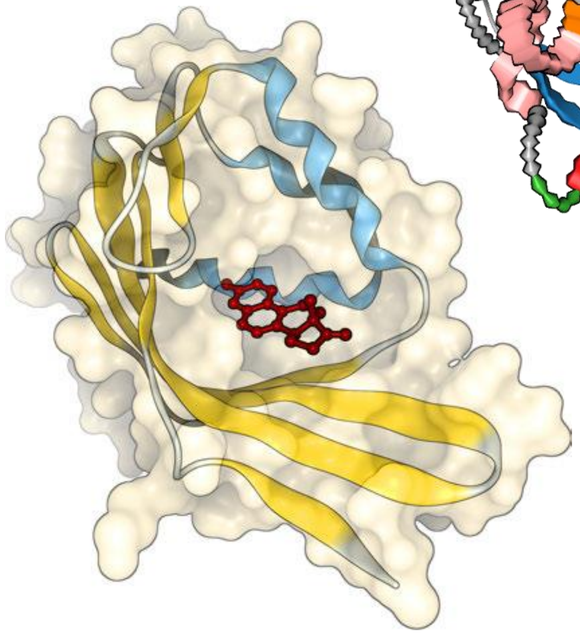
http://dendroscope.org

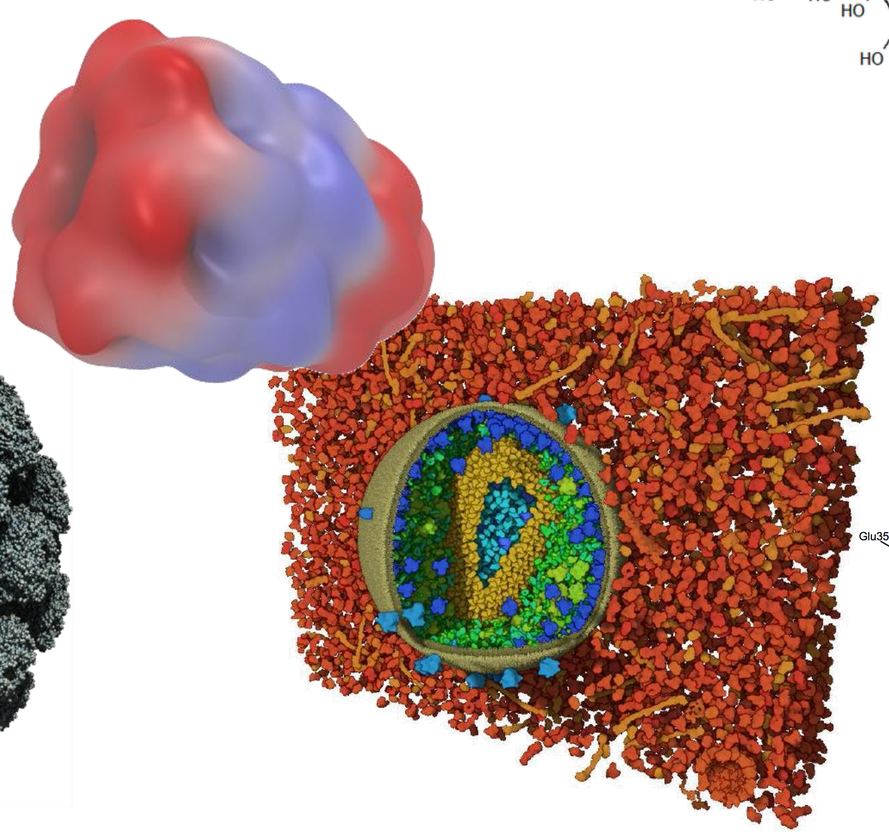
Macromolecular Structures

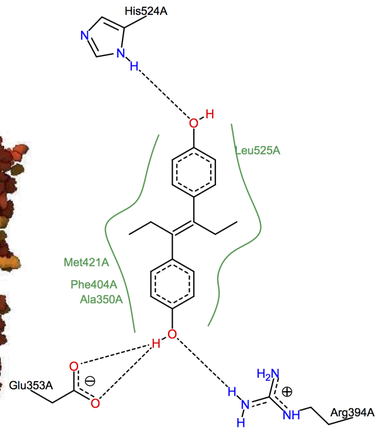




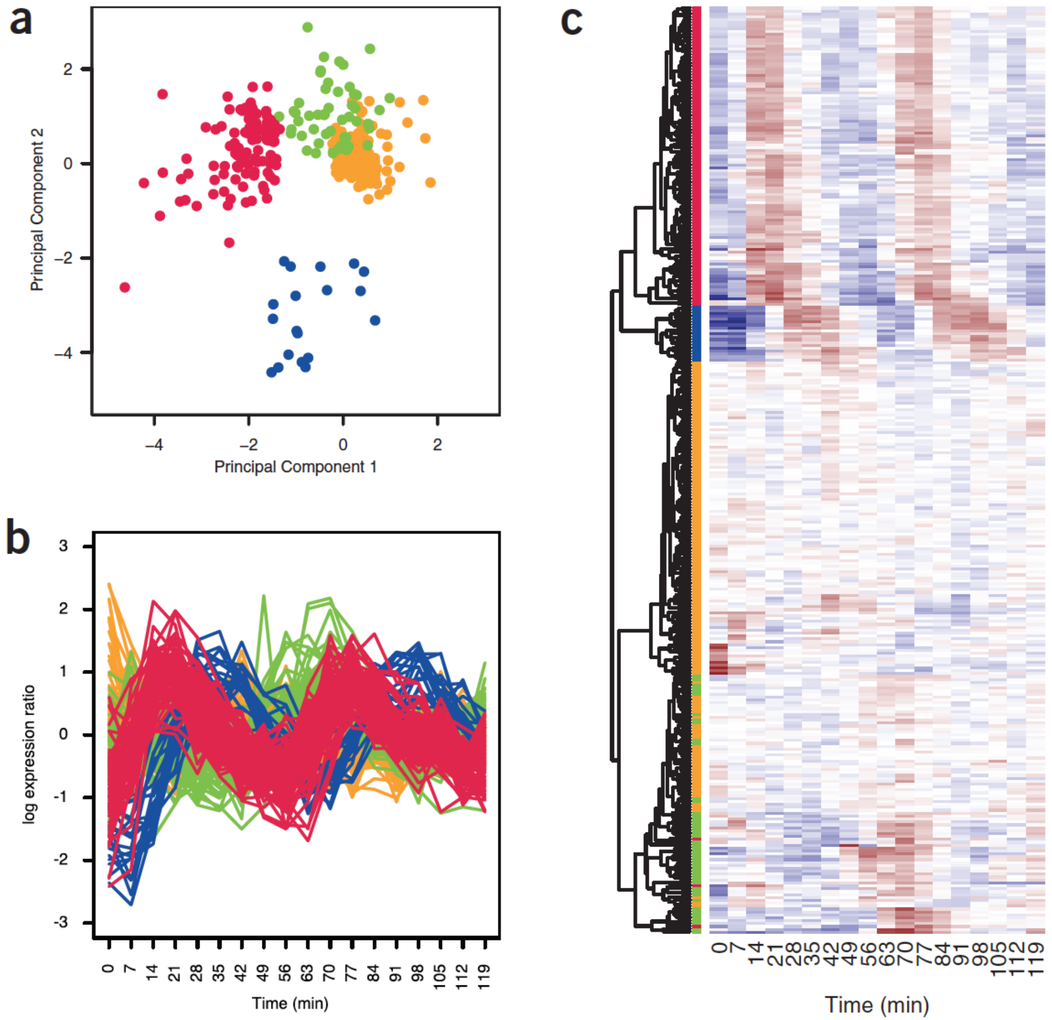






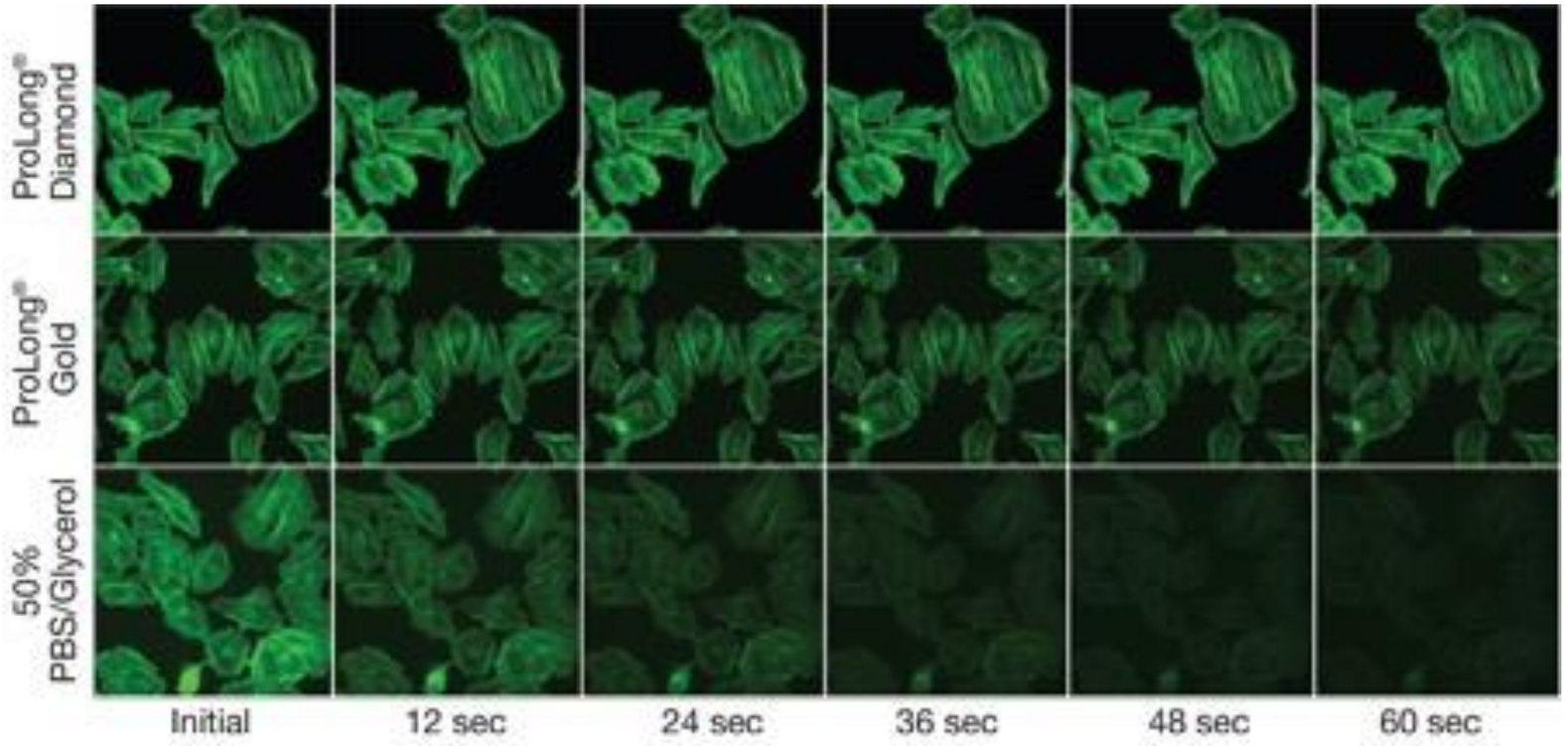


Omics Data



Gehlenborg et al. "Visualization of omics data for systems biology." Nature methods 7 (2010).

Visualization is **not** imaging!



www.thermofisher.com

Lecture topics (preliminary schedule)

- Data and task abstraction
- Design principles
- Marks and channels
- Use of color
- Plot types
- Multidimensional data
- Visual analytics
- Interactivity
- Three-dimensional visualization
- Evaluation of visualization
- Visualization in biology