

健康数据可视化

周亮

北京大学健康医疗大数据国家研究院



北京大学 健康医疗大数据国家研究院
NATIONAL INSTITUTE OF HEALTH DATA SCIENCE AT PEKING UNIVERSITY

2021-2022 第一学期

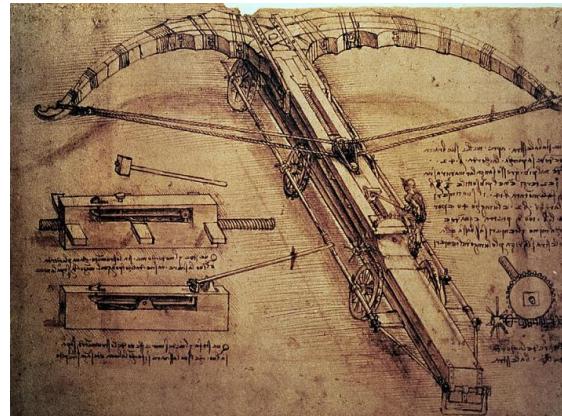
1. 可视化的意义与分类

可视化 Visualization

- 广泛意义上：构想，形成心理形象，使其可见

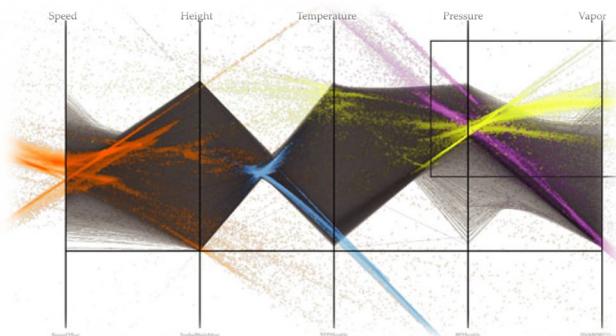


拉斯科洞穴壁画 约17000年前
https://upload.wikimedia.org/wikipedia/commons/thumb/7/7a/Lascaux-IV_26.jpg/1280px-Lascaux-IV_26.jpg

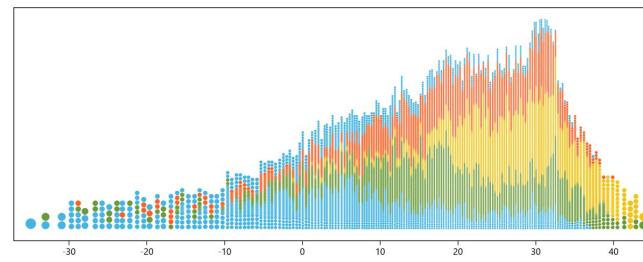


达芬奇 约15世纪
<http://www.gutenberg.org/files/20411/20411-020/20411design-for-a-giant-crossbow-refer-to-do-vinci-lb-fw=20>

- 数据科学：利用人类视觉感知和交互从数据驱动的图形图像中获取洞察——数据可视化



飓风模拟多维数据可视化
[Zhou & Weiskopf 2017, doi:
10.1109/TVCG.2017.2698041]



全球月均气温可视化
[Rodrigues & Weiskopf 2018, doi:
10.1109/TVCG.2017.2744018]



CT人体扫描可视化
[Martschink, J. et al. 2019,
doi:[10.1111/cgf.13771](https://doi.org/10.1111/cgf.13771)]

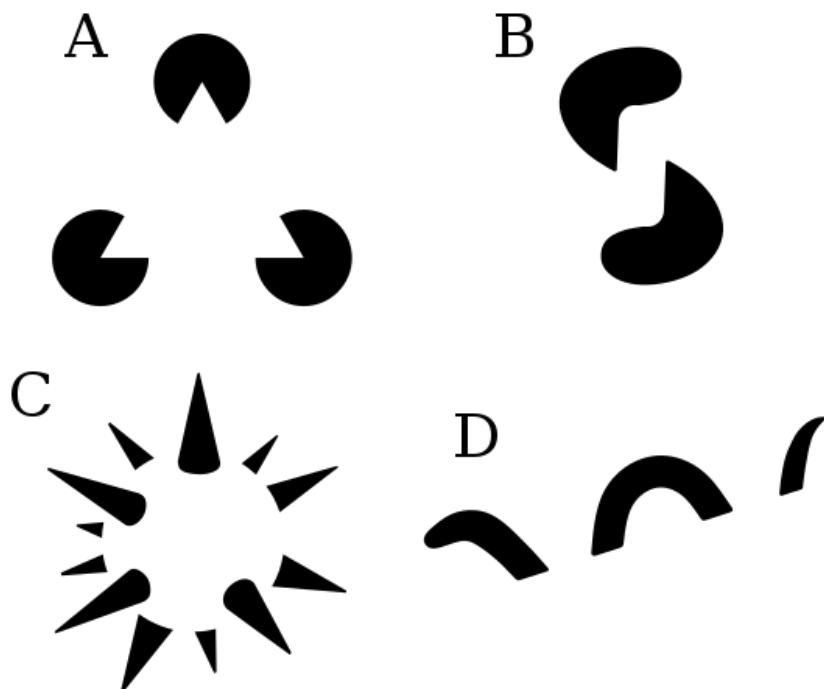
你理解的可视化?

- 你对可视化的认识?
- 日常中的可视化?
- 专业领域的可视化?

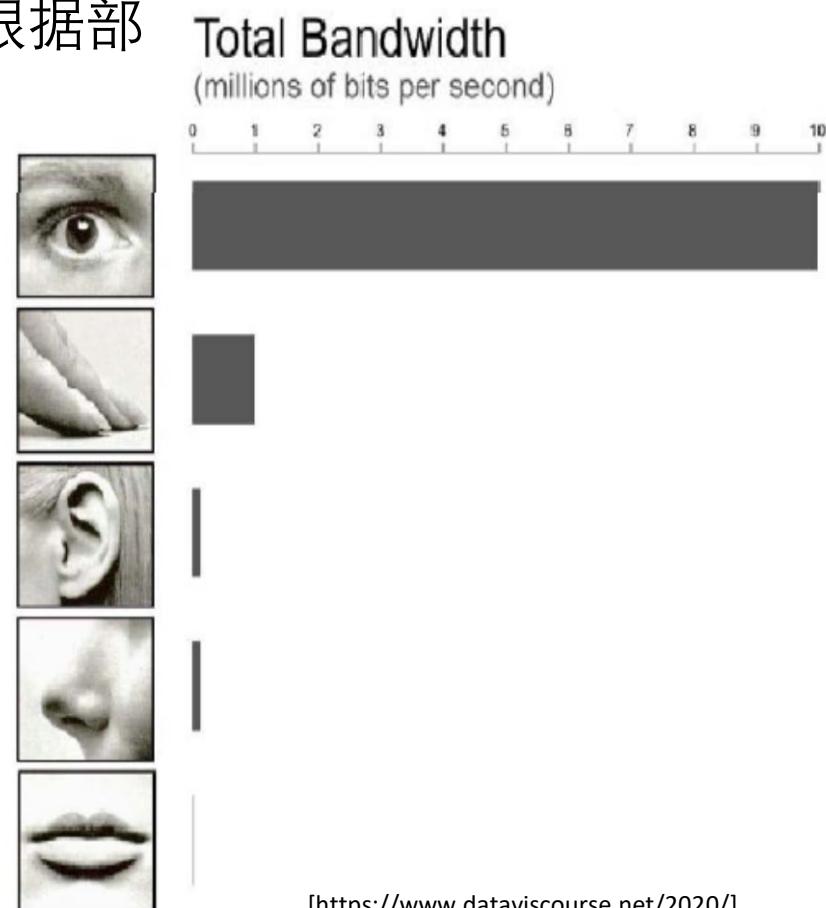


为什么要使用图形传递信息?

- 视觉是人类获取信息量最大的感知通道
- 各种视觉感知和认知现象帮助人获取整体信息
 - Gestalt (德语：造型) 理论——根据部分信息，“脑补”出整体信息



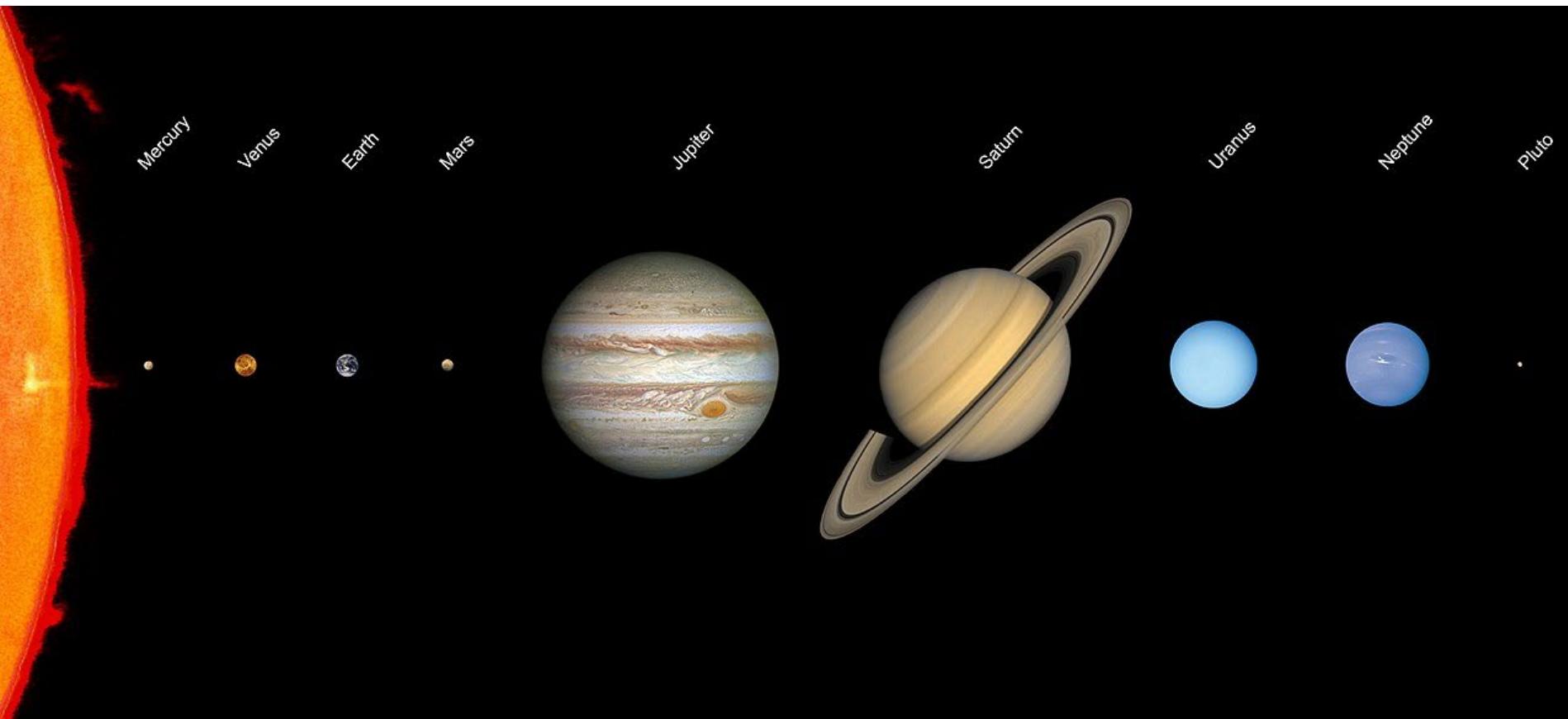
Reification [By Mrmw - Own work based on: Reification.jpg;, CC0, <https://commons.wikimedia.org/w/index.php?curid=98254077>]



[<https://www.dataviscourse.net/2020/>]

为什么要使用图形传递信息？

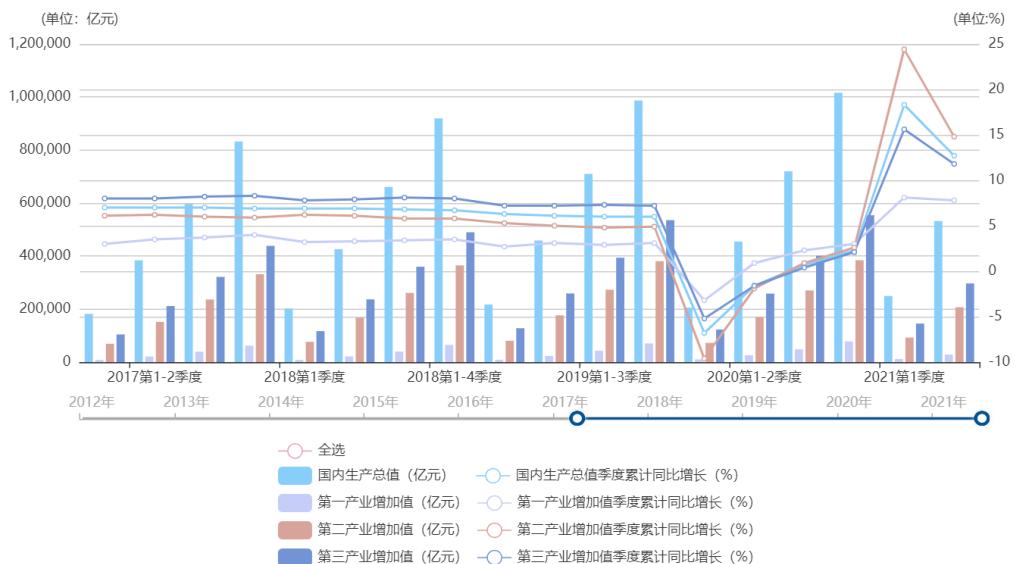
- 视觉感知带来的感受是其他感知通道无法比拟的



[By File:Solar system scale 2 wide.jpg: * File:Solar system scale.jpg: Lunar and Planetary LaboratoryFile:Jupiter and its shrunken Great Red Spot.jpg: NASA, ESA, and A. Simon (Goddard Space Flight Center)derivative work: Martin KraftFile:Sun in February (transparent).png: HalloweenNightderivative work: JCPagc2015 - This file was derived from:Solar system scale 2 wide.jpg:Sun in February (transparent).png:, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=71869737>]

可视化的意义

- 交流沟通
- 传递信息



我国国民生产总值数据
[数据详情 数据 中国政府网 (www.gov.cn)]



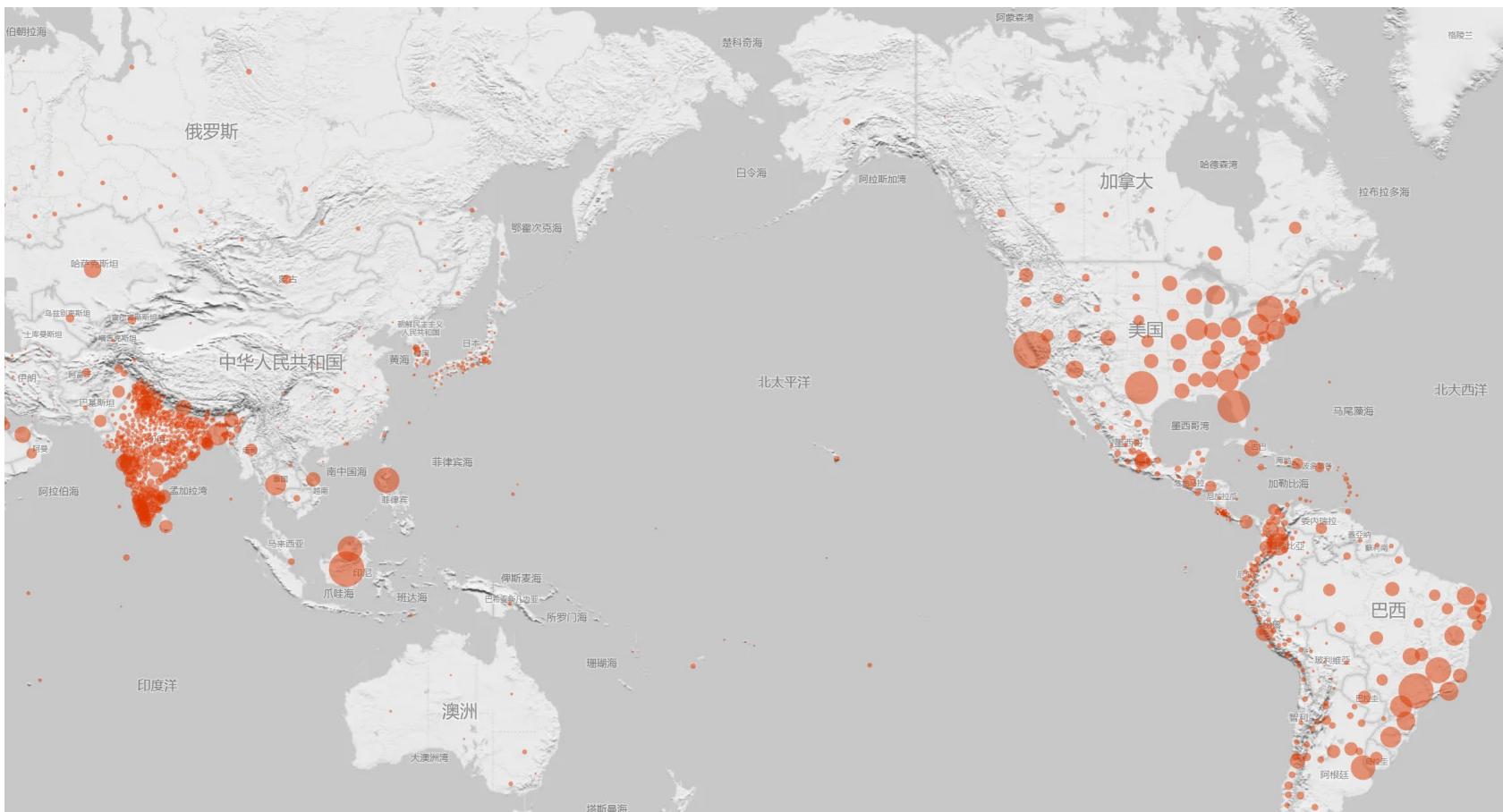
可视化与文字描述

- 文字描述：



- 数据的视觉映射带来直观对比

- 可视化：

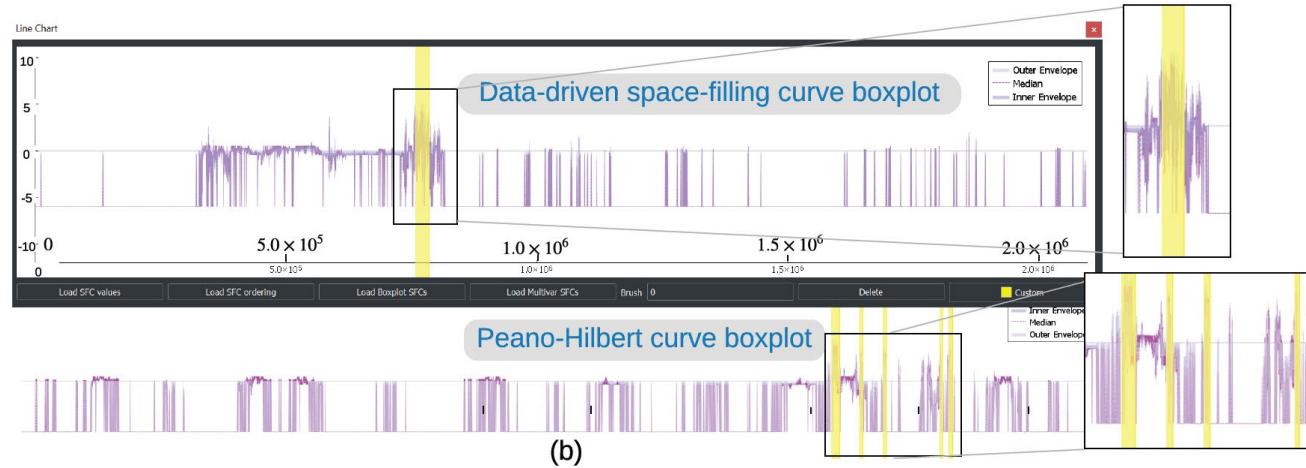


[Microsoft 必应 - 美国新冠病毒疫情实时追踪地图 (bing.com)]

©周亮 Liang Zhou

可视化的意义

- 探索
 - 找到问题的答案



在心肌缺血模拟数据集中找到电压大于3eV的空间区域?

[L. Zhou et al. (2020). Data-Driven Space-Filling Curves. doi: 10.1109/TVCG.2020.3030473.]

可视化的意义

▪ 探索——找到理想大学

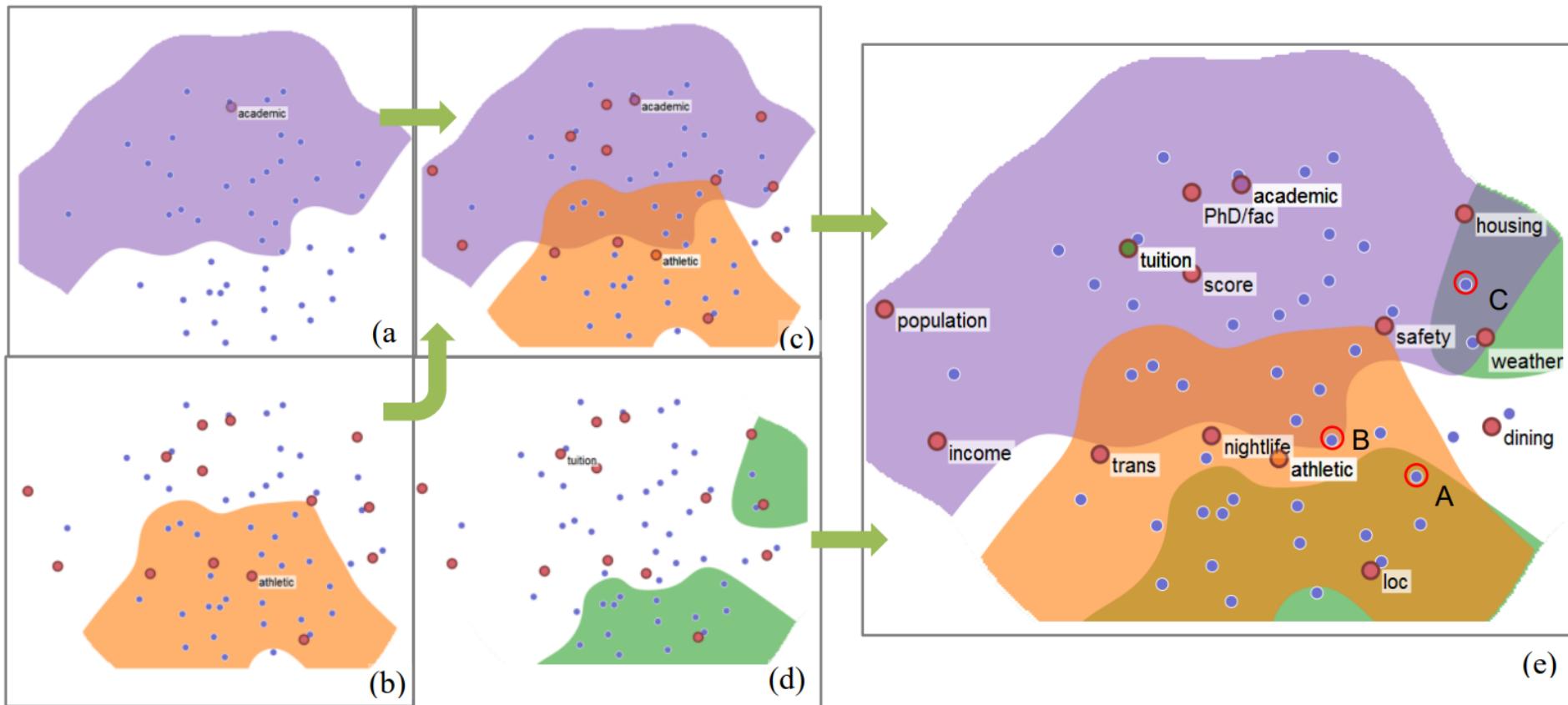
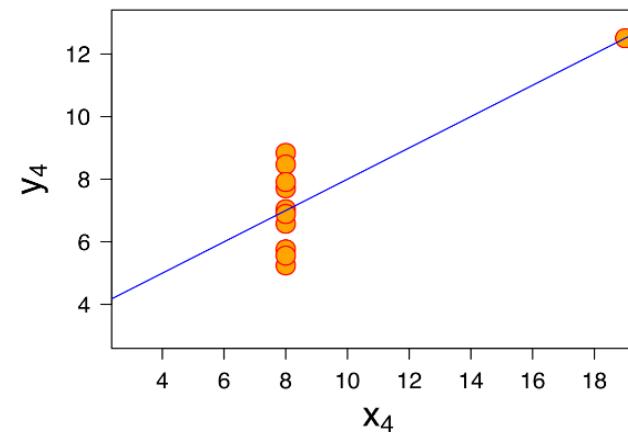
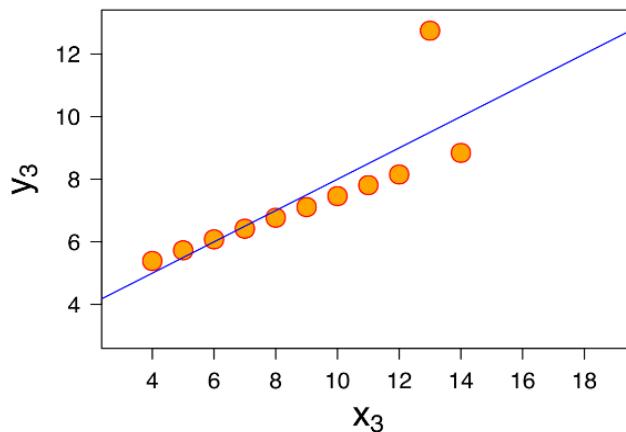
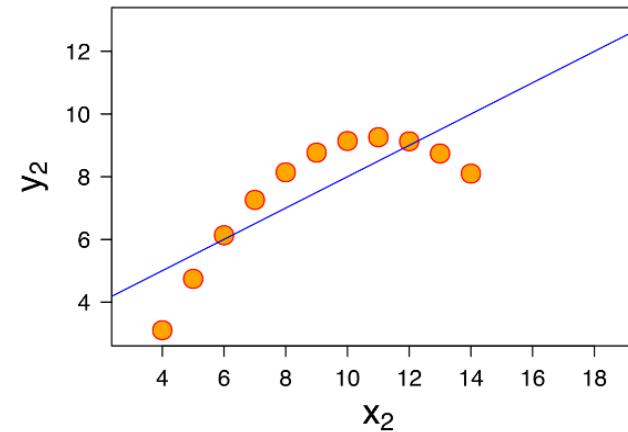
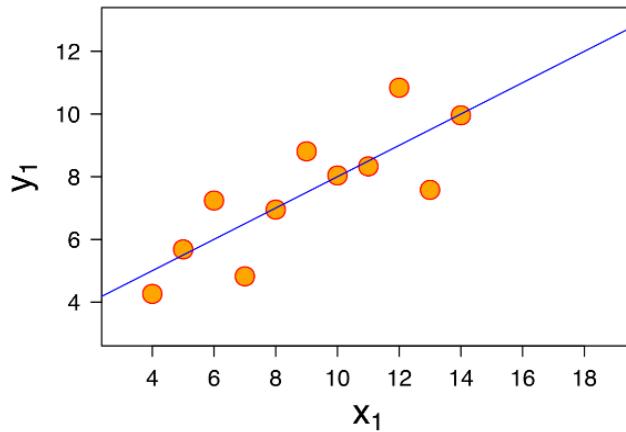


Fig.1. The process to find a dream university. (a) good academics region(>9). (b) good athletic region (>9). (c) combined region by (a) and (b). (d) low tuition region($<\$18,000$). (e) combined region by region (c) and (d).

[S. Cheng and K. Mueller, "The Data Context Map: Fusing Data and Attributes into a Unified Display," doi: 10.1109/TVCG.2015.2467552.]

可视化与统计

- 统计结果完全相同的数据可能完全不同
- Anscombe's Quartet

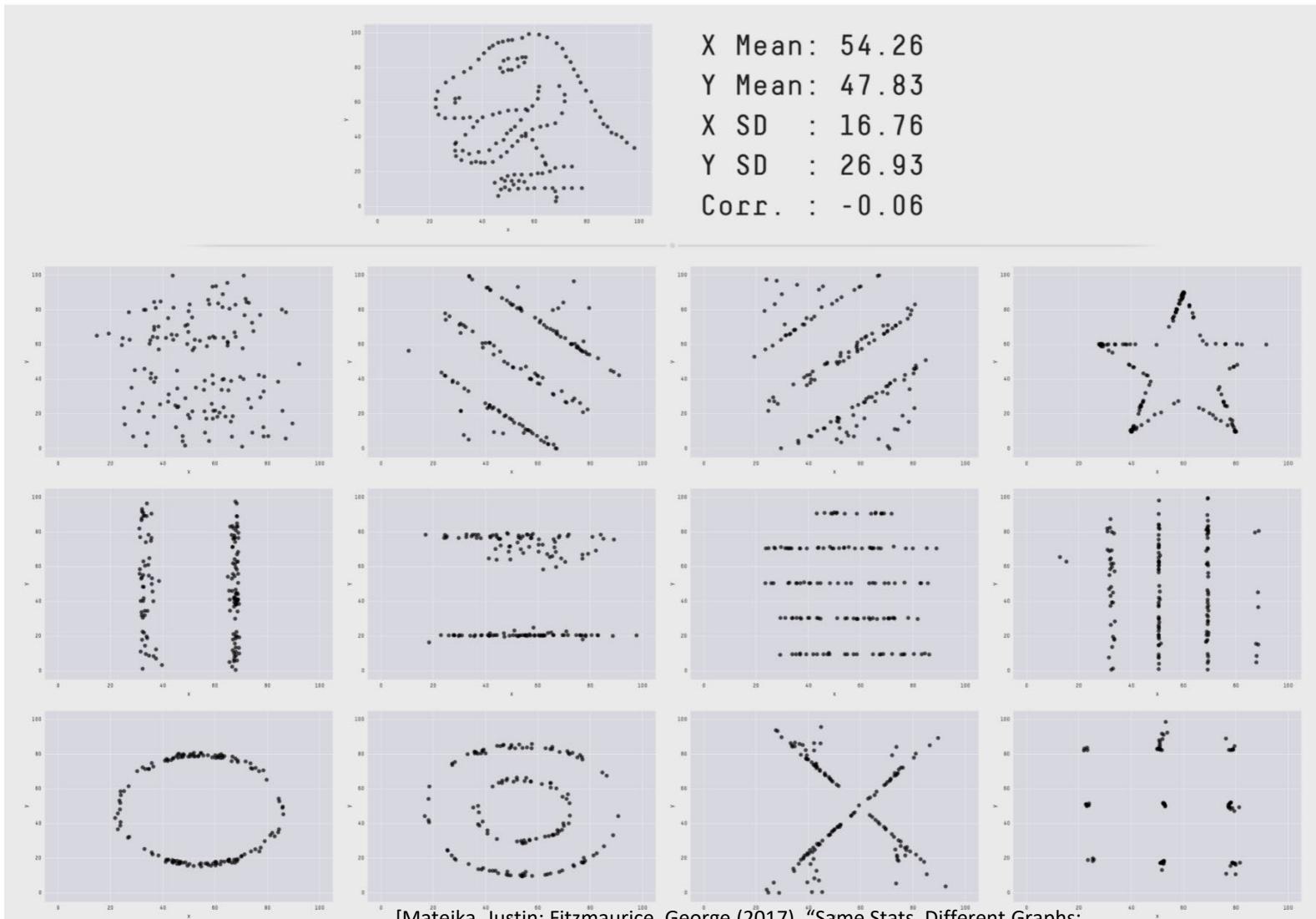


[By Anscombe.svg: Schutz (label using subscripts): Avenue - Anscombe.svg, CC BY-SA 3.0.]

©周亮 Liang Zhou

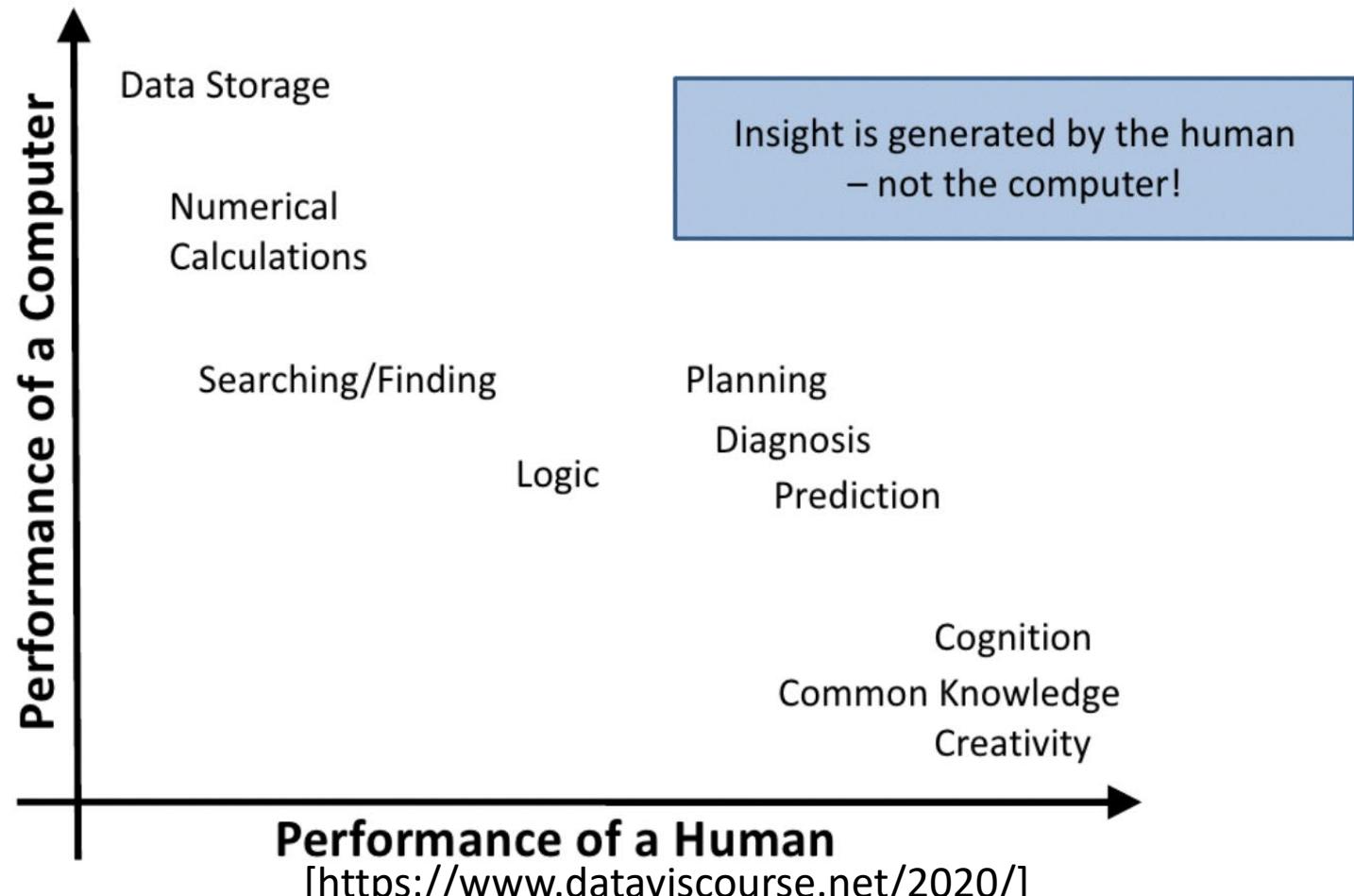
可视化与统计

- “..make both calculations and graphs. Both sorts of output should be studied; each will contribute to understanding.” – F.J. Anscombe



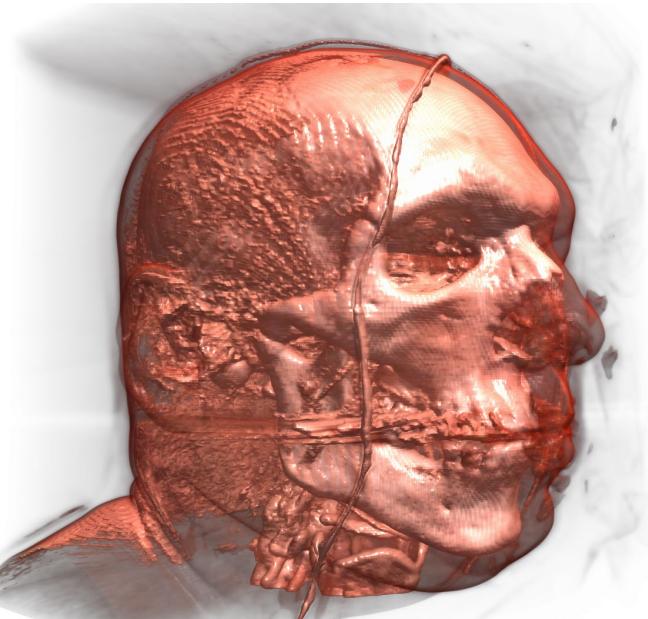
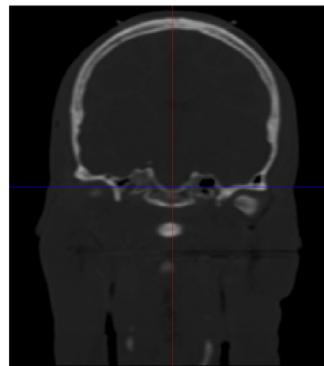
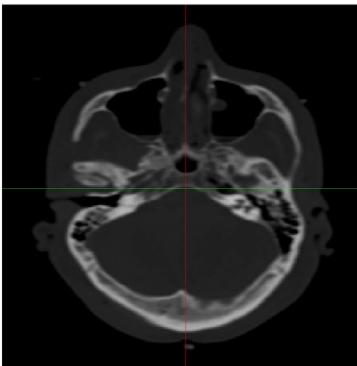
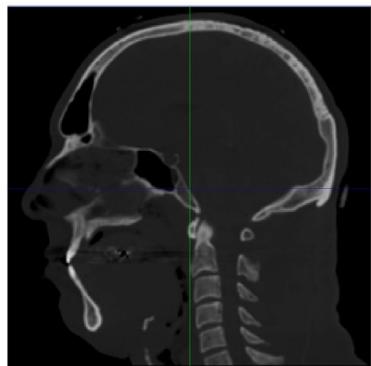
可视化与计算机自动

- 既然有了自动化方法，为什么还需要可视化？
- 计算机擅长计算、存储、搜索，解决标准化数据上良定义的问题
- 计算机无法生成洞察，不能进行有创造力的工作



可视化与计算机自动

- 各取所长→计算机计算存储，人交互探索
- 例子：三维空间数据可视化
 - 如何手工绘制医学影像的三维可视化？

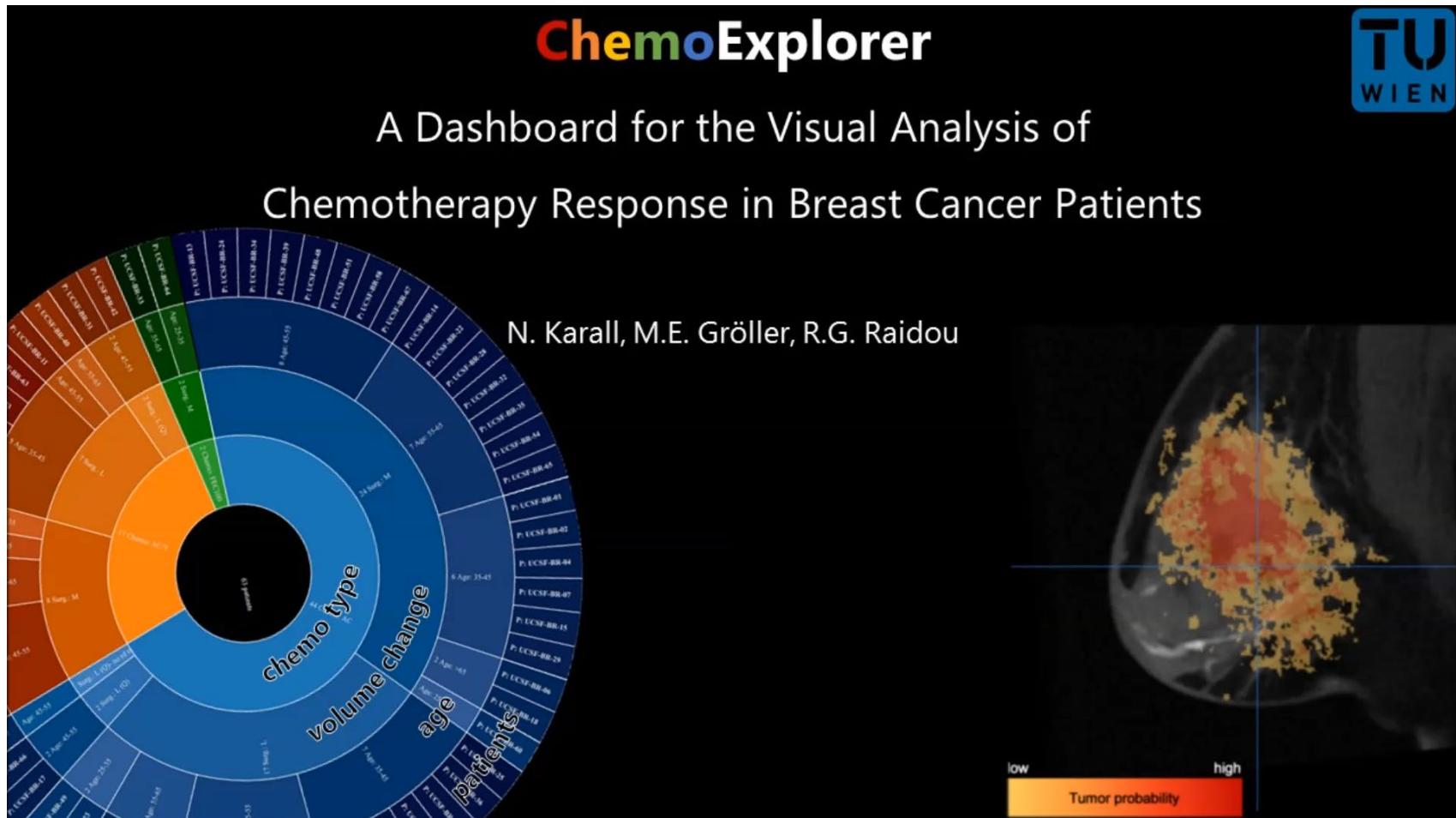


CT数据来自Visible human
project: male



可视化与计算机自动

- 计算机的计算能力支持人进行交互
 - 三维数据可视化例子，即使可以手绘，也无法绘制所有角度的可视化，进行交互。
- 支持人进行交互→使人在探索、决策中发挥主导作用



[N. Karall, et al. "ChemoExplorer: A Dashboard for the Visual Analysis of Chemotherapy Response in Breast Cancer Patients.", doi:10.2312/eurovisshort.20181077]

©周亮 Liang Zhou

可视化的要素

- 可视化=人+数据+视觉映射+交互

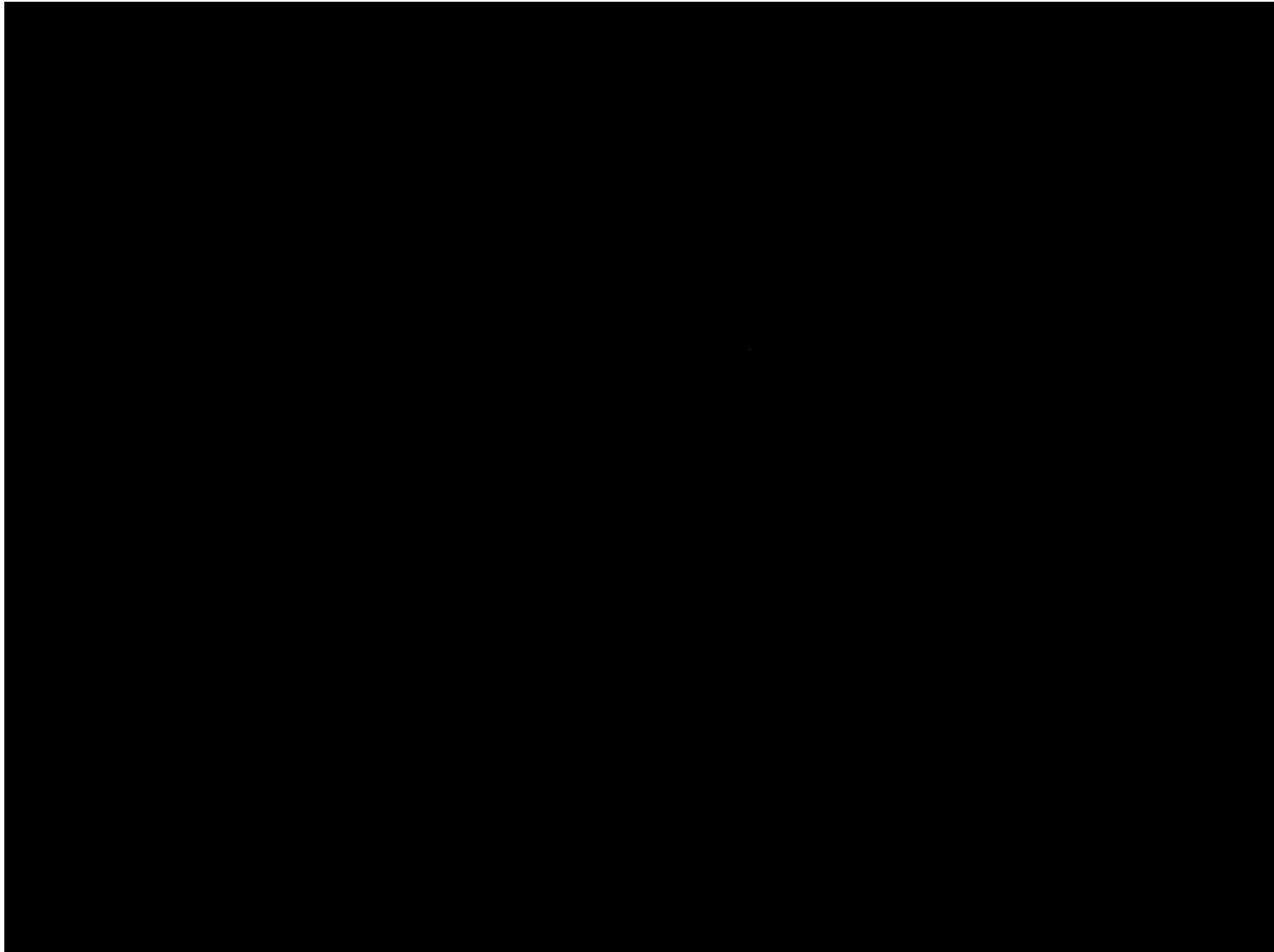
Visualization =

Human Data Interaction
+ Visual Mapping

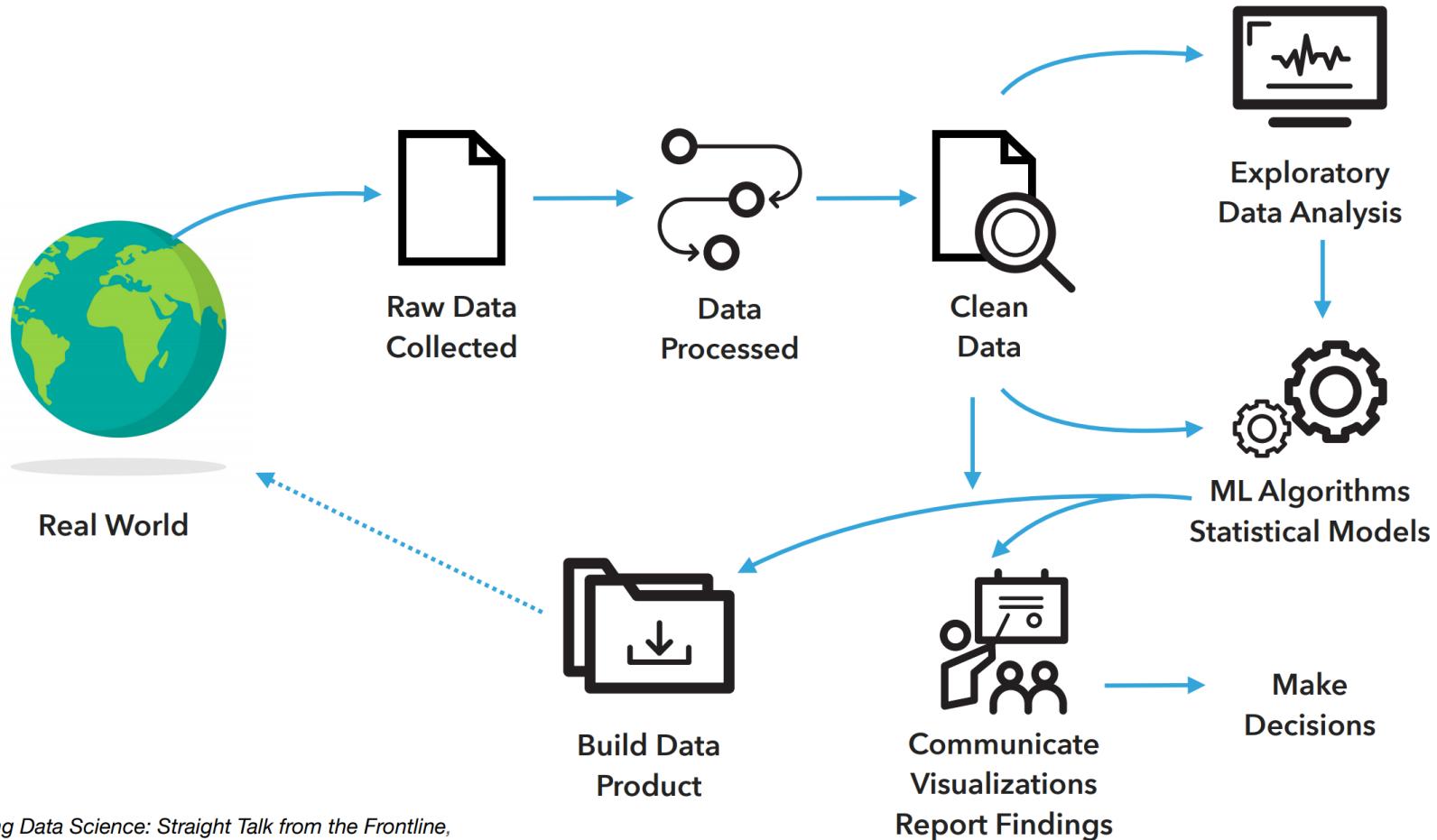
[<https://www.dataviscourse.net/2020/>]

人机交互

- 计算机的重要研究和应用领域



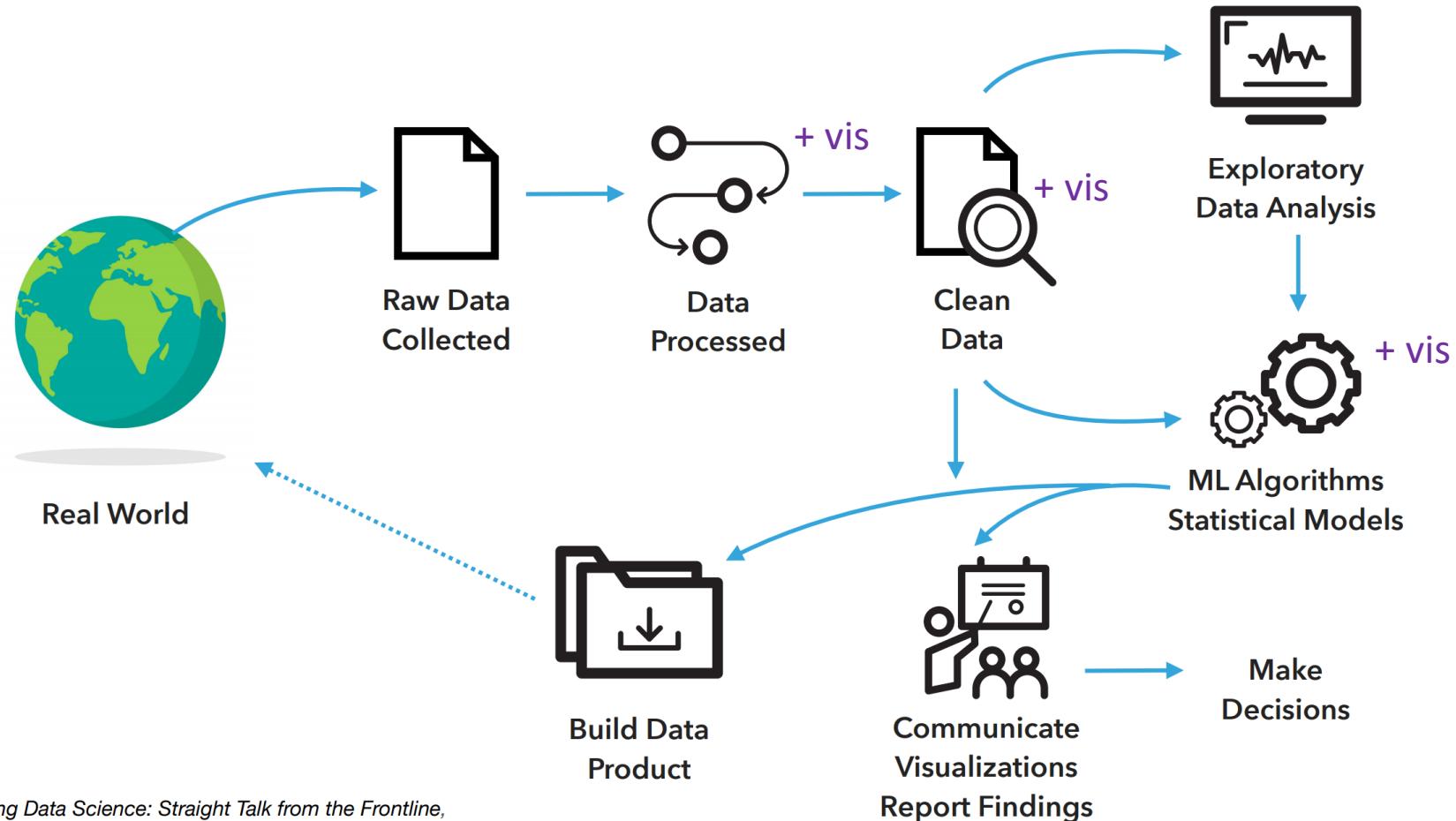
可视化与数据科学



Adapted from *Doing Data Science: Straight Talk from the Frontline*,
[O'Neil and Schutt, 2013]

可视化与数据科学

- 可视化能参与数据科学中各步骤



Adapted from *Doing Data Science: Straight Talk from the Frontline*,
[O'Neil and Schutt, 2013]

历史上的可视化

▪ 新刊铜人针灸经，七卷

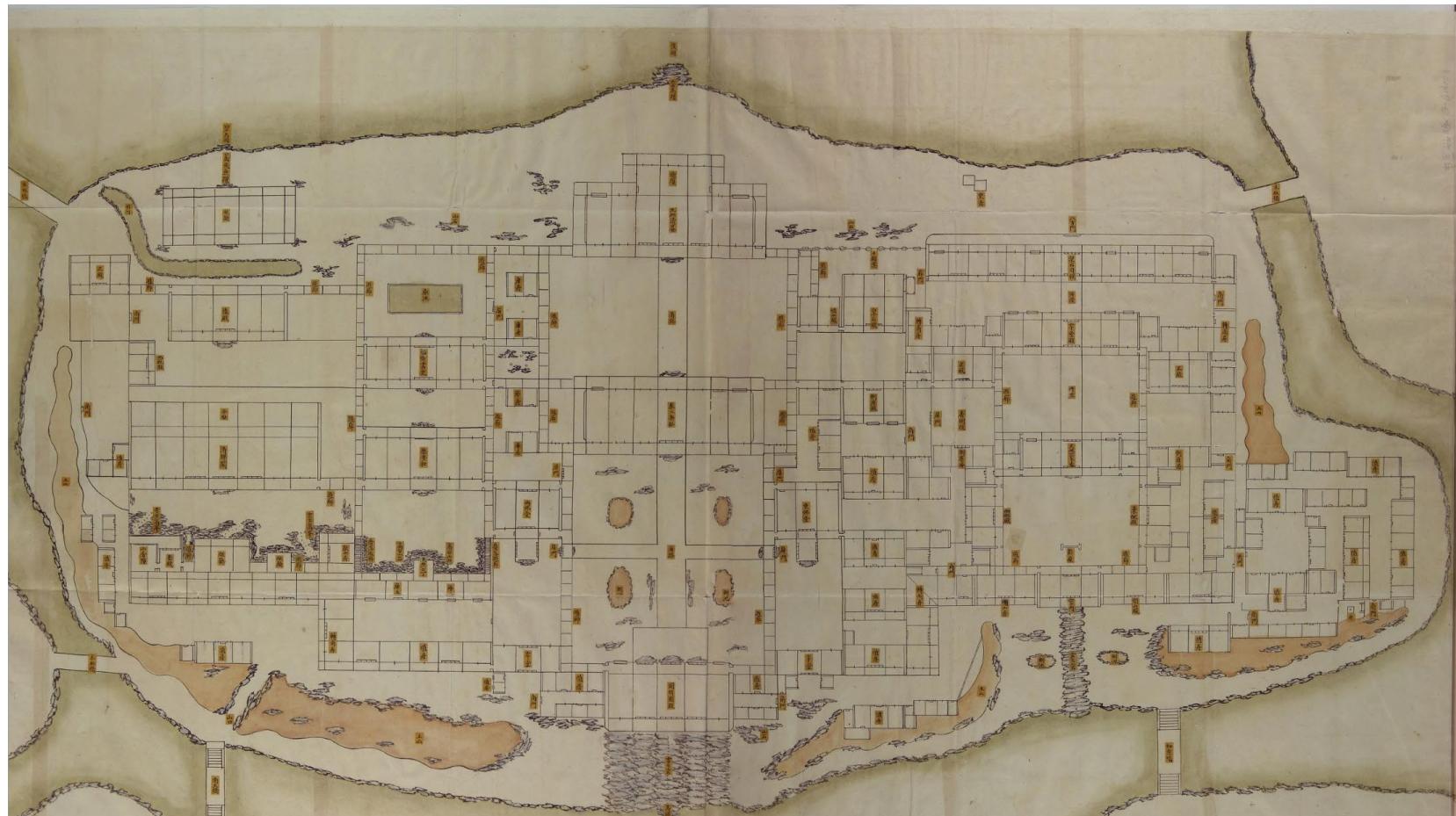


[中国国家图书馆·新刊铜人针灸经，七卷. <https://www.wdl.org/zh/item/13512/>]

©周亮 Liang Zhou

历史上的可视化

▪ 样式雷图档



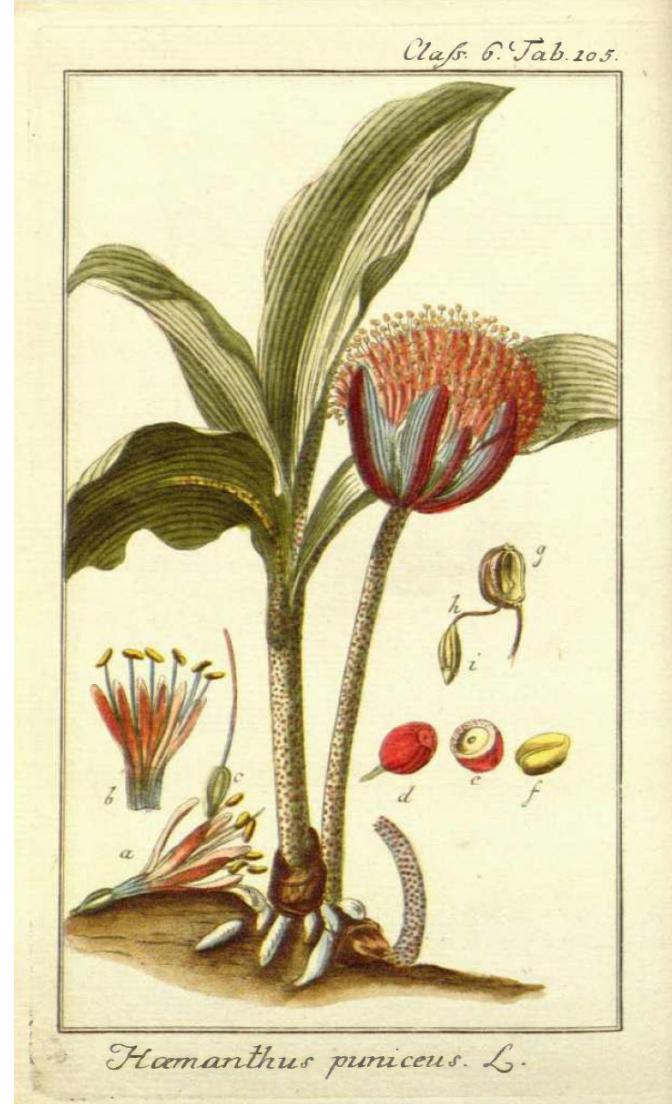
[中国国家图书馆. 样式雷图档 (3) 颐和园排云殿佛香阁地盘图.]

历史上的可视化

▪ 科学插图



[By Maria Sibylla Merian - from Metamorphosis insectorum Surinamensis, Plate XLIII., Public Domain, <https://commons.wikimedia.org/w/index.php?curid=8148639>]



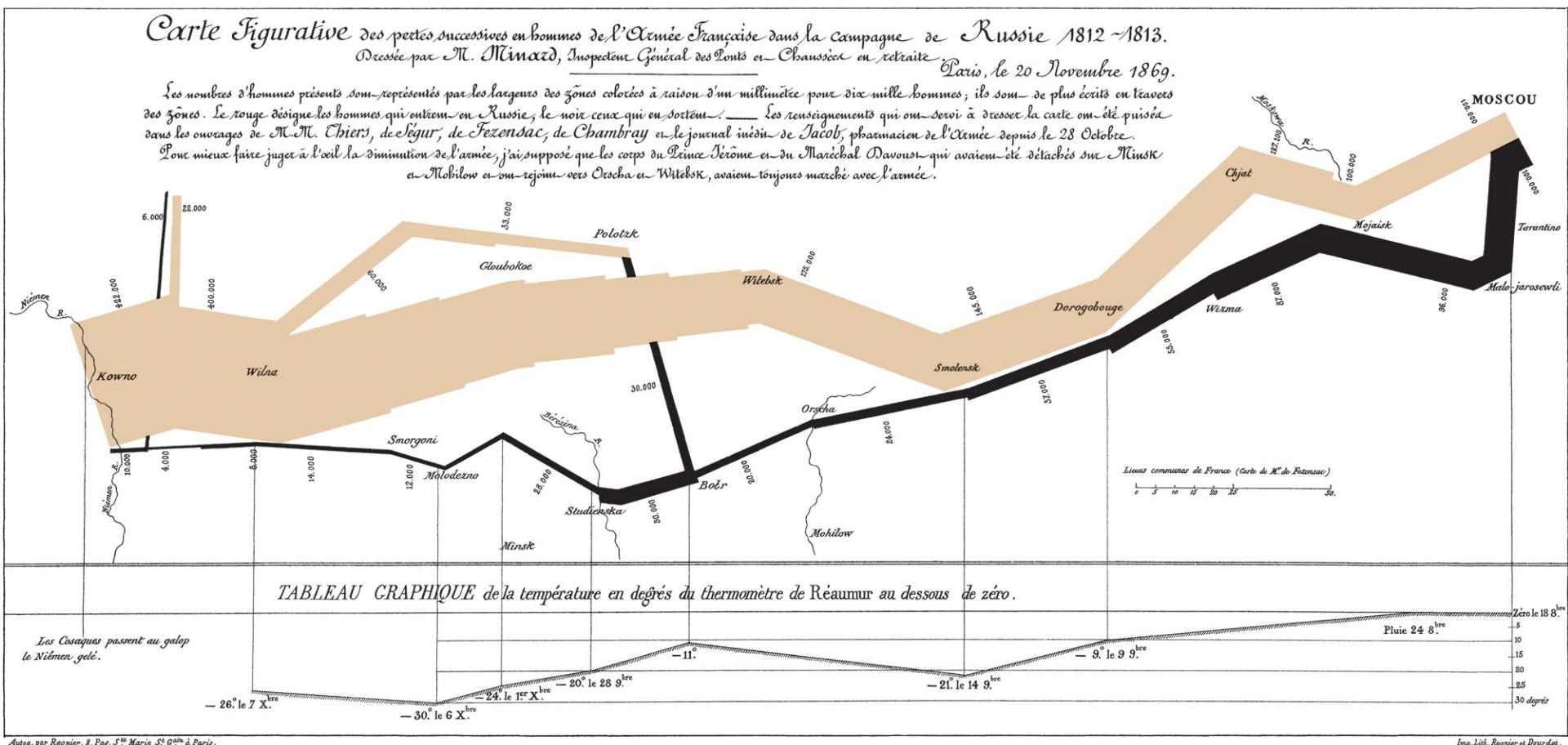
[By Zorn, Johannes, 1739-1799. Jacquin, Nikolaus Joseph, Freiherr von, 1727-1817 - <http://www.botanicus.org/title/b11954188>, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=19358489>]

©周亮 Liang Zhou

历史上的可视化

▪ 数据分析

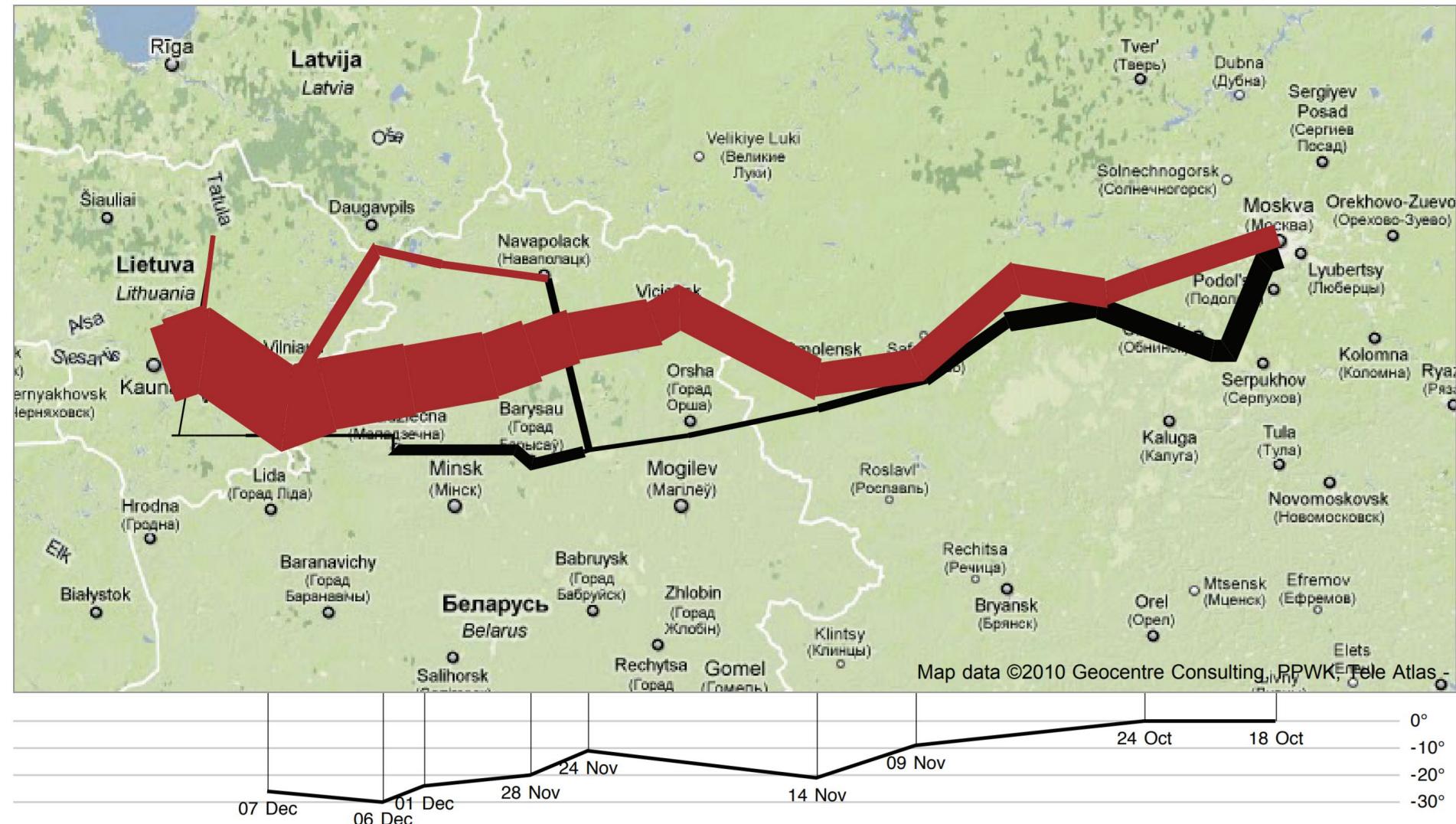
▪ Charles Minard 拿破仑俄国战役 (1812-1813) 损失地图



[By Charles Minard (1781-1870) - see upload log, Public Domain,
<https://commons.wikimedia.org/w/index.php?curid=297925>]

拿破仑俄国战役地图

- 空间+时间+军队人数+气温

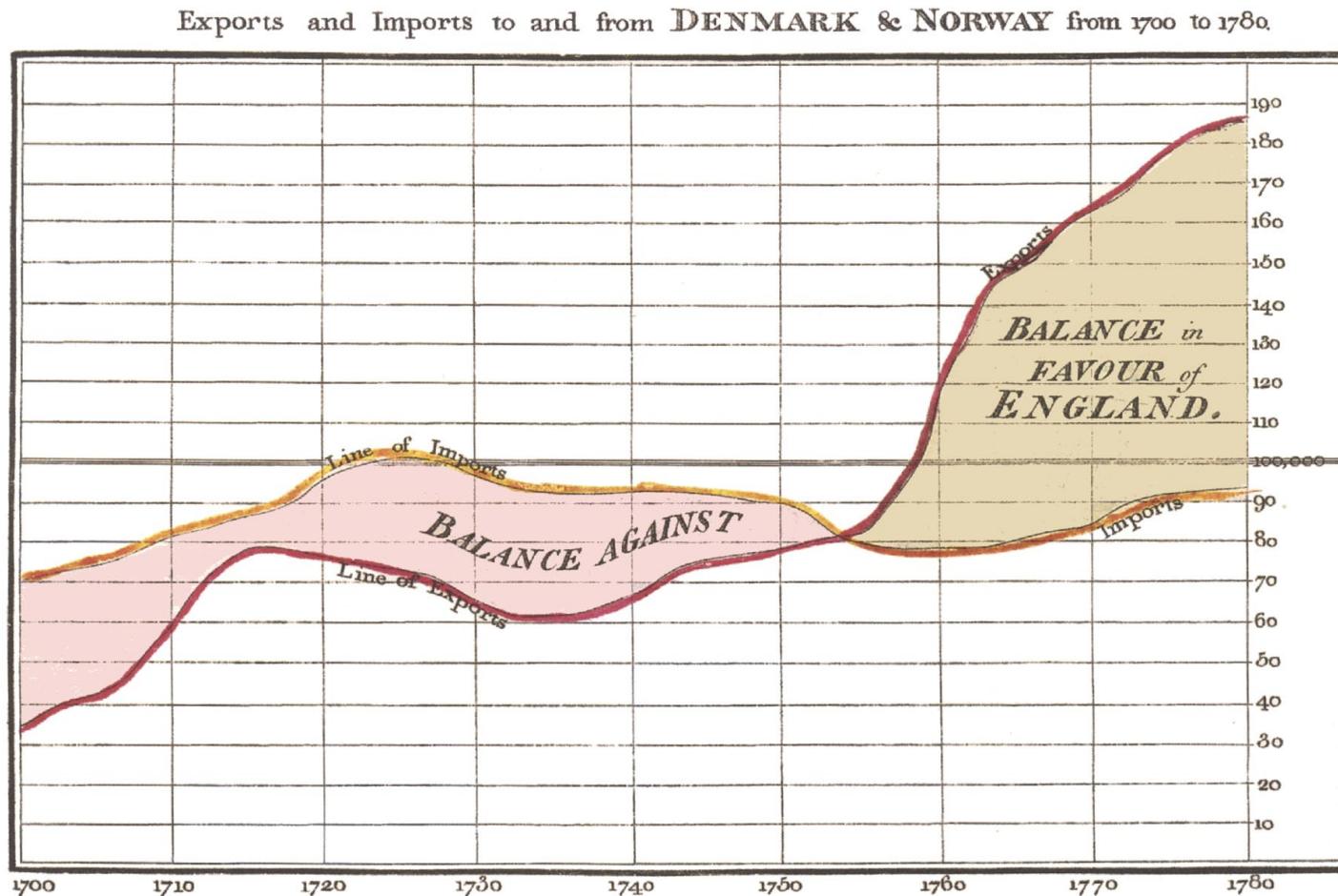


计算机重建[<https://homes.cs.washington.edu/~jheer//files/zoo/ex/maps/napoleon.html>]

©周亮 Liang Zhou

历史上的可视化

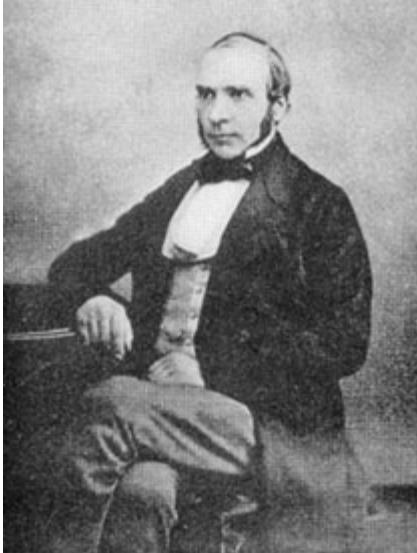
▪ 数据分析——贸易的收支时序分析



历史上的医学可视化

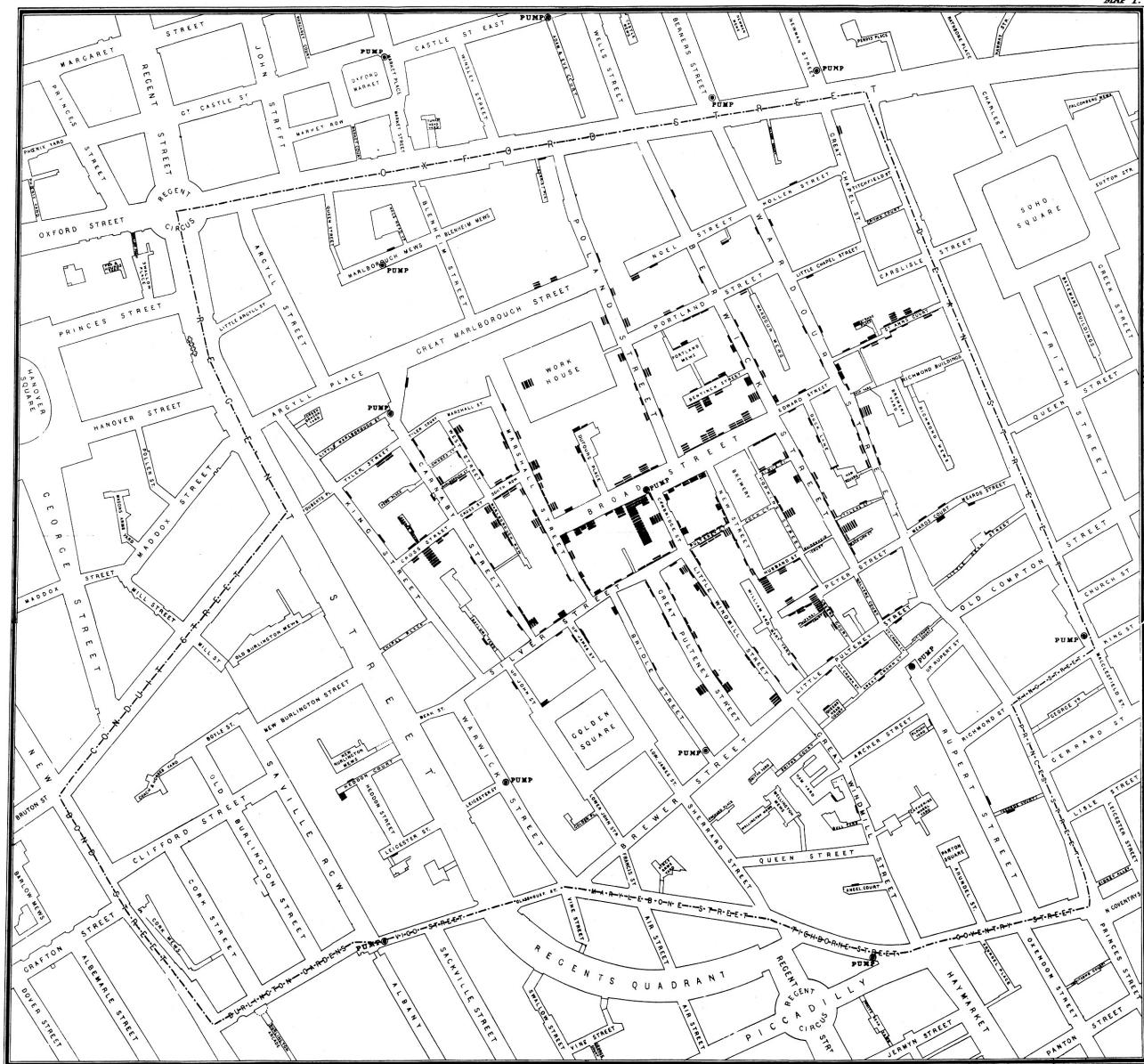
- Cholera map 1854 --- J. Snow

- ## ▪ 霍乱传染模式研究



[Originally from en.Wikipedia, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=403227>]

[By John Snow - Map of the book; On the Mode of Communication of Cholera; by John Snow, originally published in 1854. Digitally enhanced version found on the UCLA Department of Epidemiology website., Public Domain, <https://commons.wikimedia.org/w/index.php?curid=2278605>]



历史上的医学可视化

▪ 霍乱传染模式——水源污染



John Snow, 1854

E.Tufte, Visual Explanations, 1997



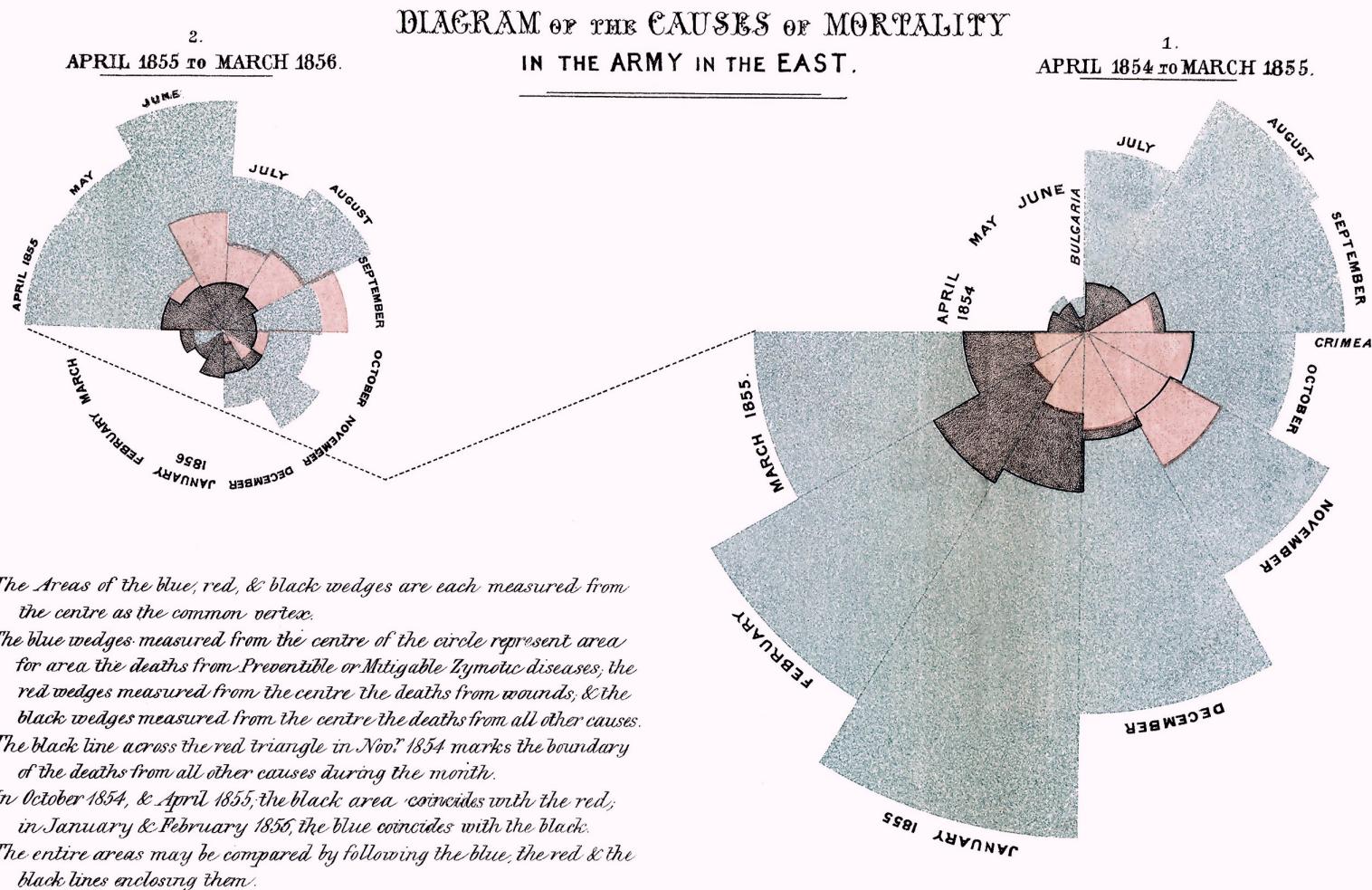
历史上的医学可视化

- Rose chart 1859 --- F. Nightingale
- 克里米亚战争死亡原因分析

[By w:Florence Nightingale (1820–1910). - Public Domain, <https://commons.wikimedia.org/w/index.php?curid=1474443>]

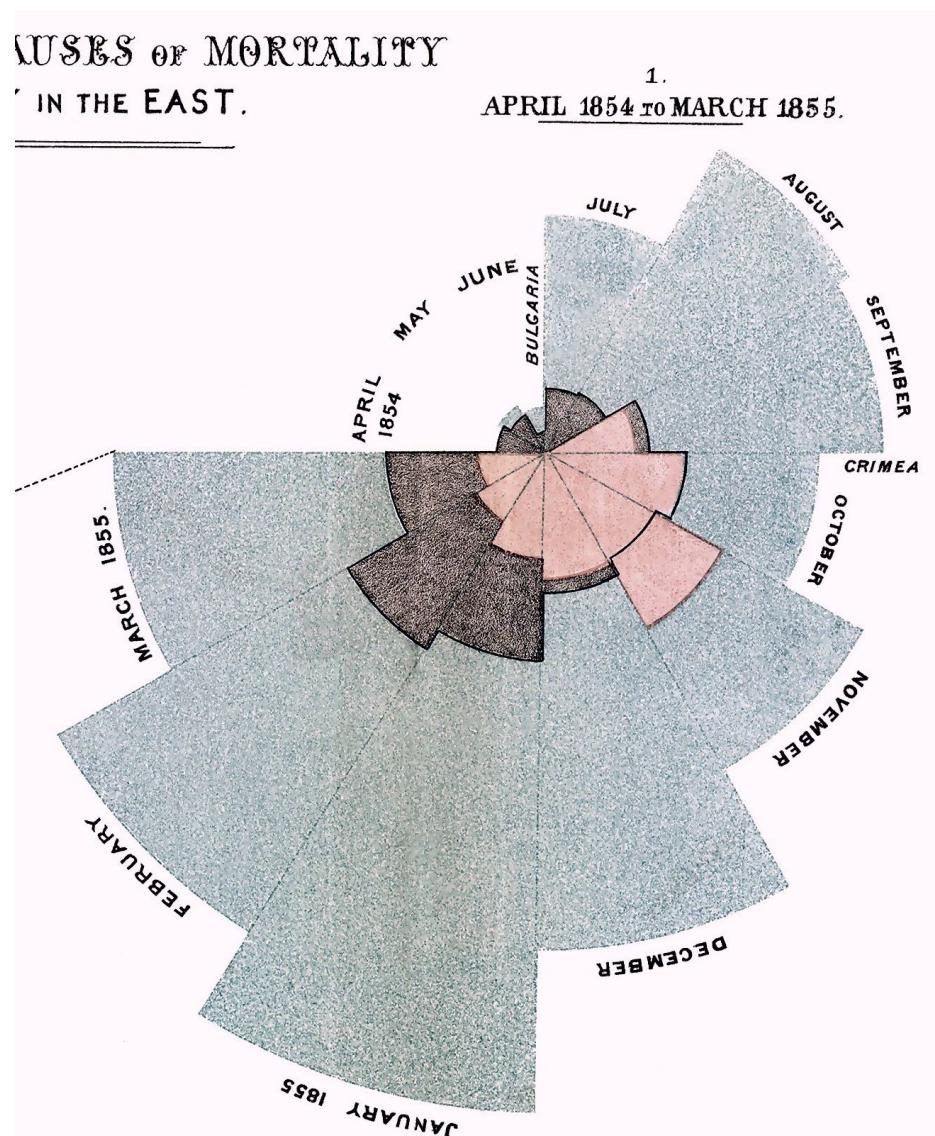


[By Henry Hering (1814-1893) - National Portrait Gallery, London, Public Domain, <https://commons.wikimedia.org/w/index.php?curid=61520366>]



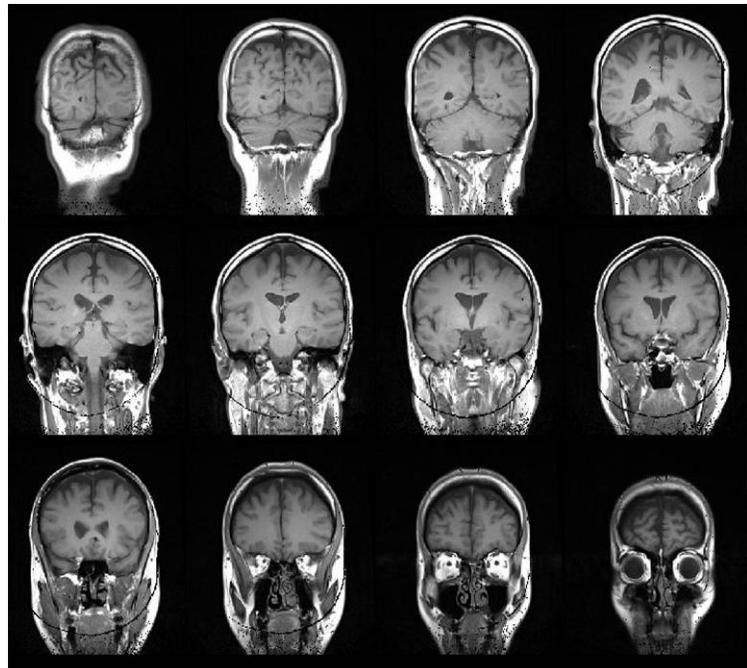
历史上的医学可视化

- 死亡原因分析——最大比例为可预防的疾病



可视化与健康医疗

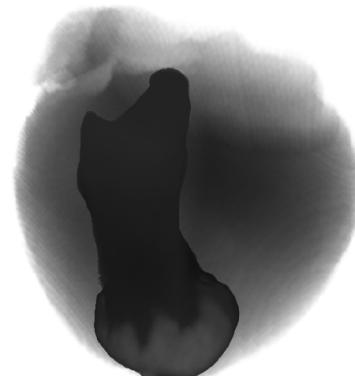
- 对健康医疗数据进行可视化、可视分析
 - 提高健康医疗数据分析能力；辅助健康医疗决策
- 辅助健康医疗过程
 - 提高健康医疗实践水平



脑部MRI数据

http://4.bp.blogspot.com/-BPNZ9LzfX4/TqFoHX5dOSI/AAAAAAAFAU/F316MyCLHZo/s1600/pre_nogad_montage34_edit.jpg

心肌缺血模拟数据体渲染
[Source: L. Zhou]



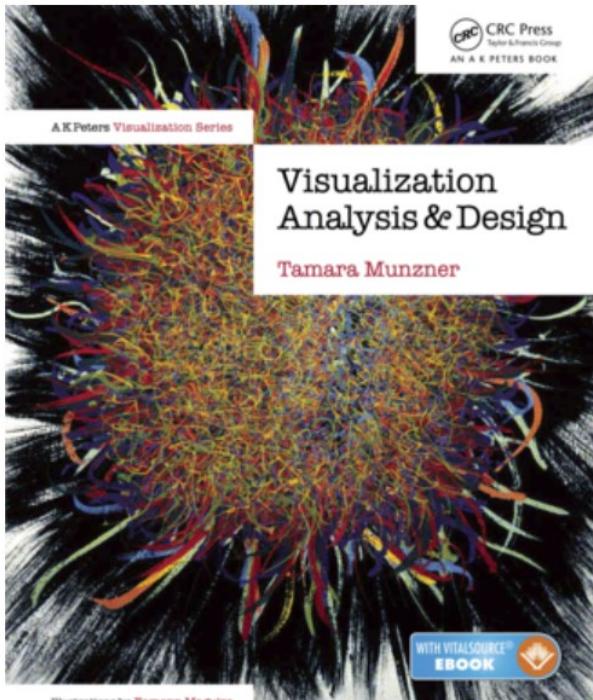
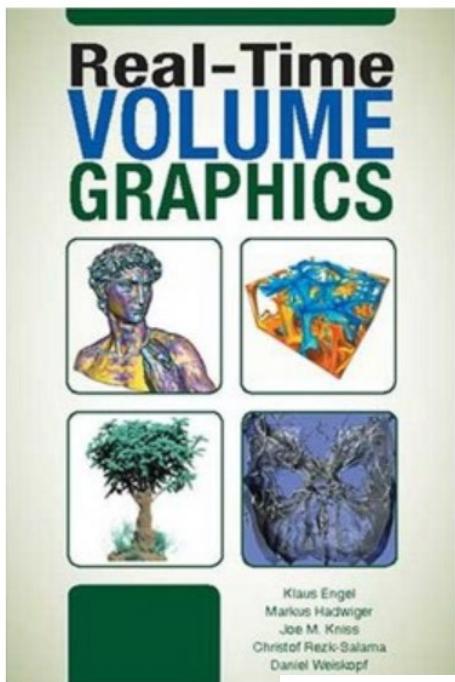
增强现实辅助手术

<https://immersive-technology-healthcare.hcs-pharma.com/2019/03/01/augmented-reality-surgical-technology-unveiled-by-philips-and-microsoft/>

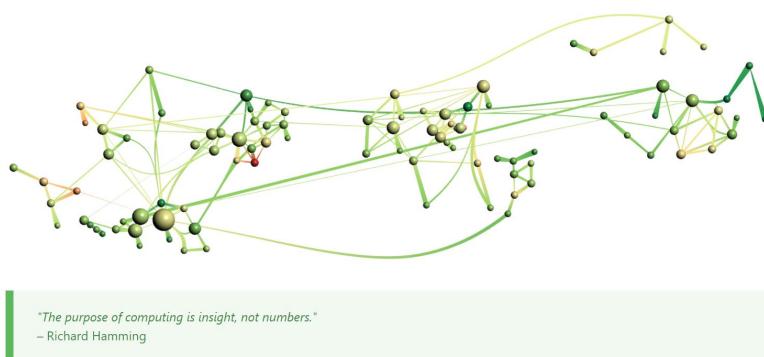
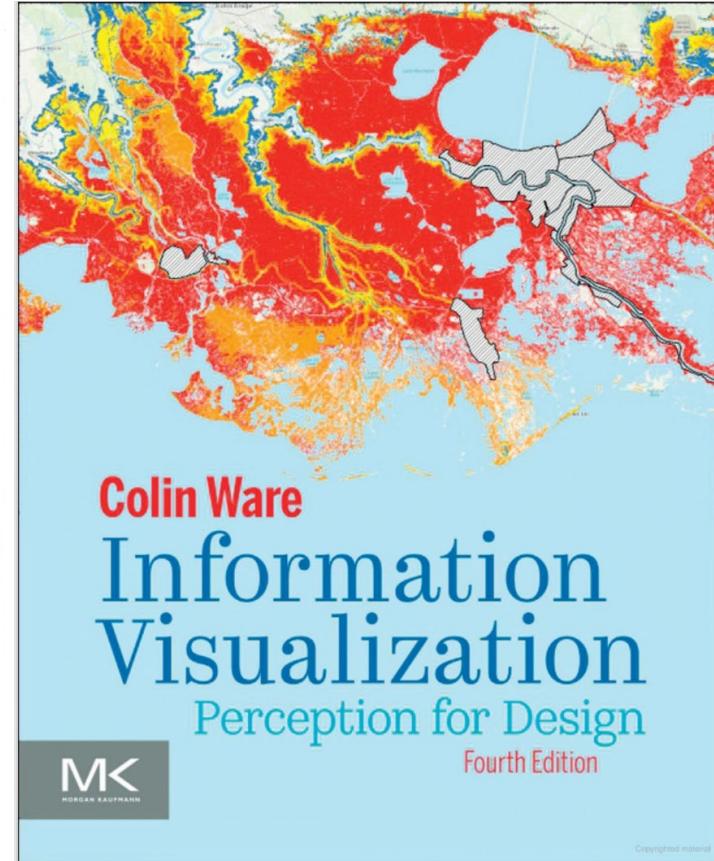
本课程目标

- 能看懂可视化，建立视觉思考能力
- 了解可视化的基本方法
- 能进行简单的可视化编程
- 对指定数据，了解适用的可视化方法
- 能评价可视化方法的优缺点
- **建立健全医疗和可视化的桥梁，跨学科沟通**
- **<https://zhou-l.github.io/>**

推荐书目



Interactive Visual Data Analysis
Christian Tominski and Heidrun Schumann
AK Peters Visualization Series
CRC Press



可视化处理数据种类差异巨大

- 影像数据（二维，三维）

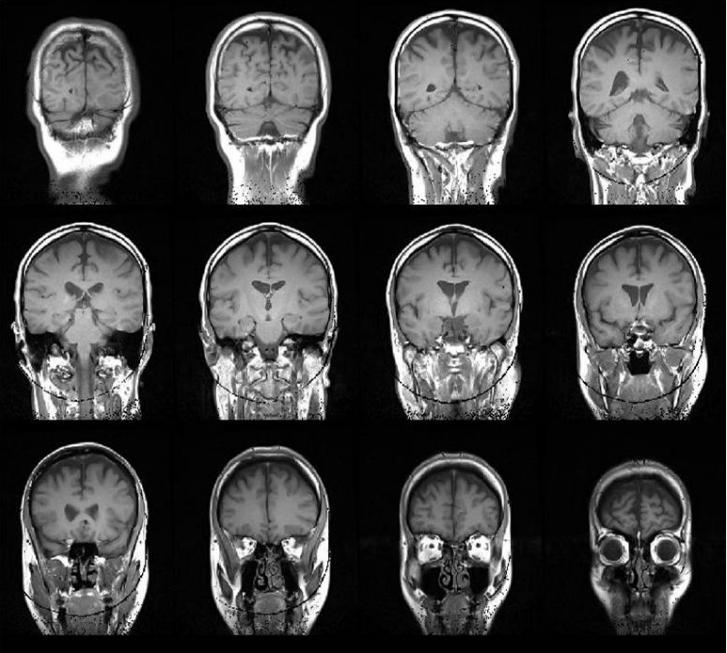
- 表格数据

- 层次结构数据

- 非结构数据

-

1	Symbol	N11	N14	N1	N27	N3	N8	CDJIE	DHZ	JMY	LMENG	L5L	SGH	WHRU	WSY
2	A1BG	0.5645	0.1503	0.4451	0.1492	0.7234	0.1117	0.0645	0.3301	0.4823	0.0997	0.4797	0.4754	0.4303	0.3398
3	A2M	197.0306	93.9088	90.2652	360.0319	273.0878	80.8098	79.7098	166.174	101.5462	87.3495	104.0838	172.3037	140.6927	106.8786
4	A2ML1	26.9459	6.734	49.7505	13.1724	33.4317	43.8277	30.9469	4.6351	17.5137	24.9785	23.1412	6.1723	16.9964	11.1661
5	A2MP1	0.2883	0	0	0	0.031	0	0.4162	0.8802	0.4702	0.3145	0.748	1.1724	1.234	0.6902
6	A3GALT2	0	0	0.091	0	0	0	0.1263	0.2585	0.2518	0	0.1174	0.2482	0	C
7	A4GALT	26.5808	13.3641	12.6768	29.3892	39.2246	13.3254	21.2073	37.2913	11.0729	9.9988	18.2436	31.4095	17.3294	21.7379
8	AAAS	38.2259	37.178	29.1431	32.3595	37.8547	29.8753	21.0786	23.7101	16.8398	10.9726	15.3844	18.1043	16.6646	19.3478
9	AACS	45.2092	56.787	92.5636	31.6176	29.709	87.9038	21.6945	9.933	43.1694	42.0701	19.6508	34.2724	43.0972	19.8138
10	AACSP1	0.5753	0.2452	0.3629	0.1217	0.4537	0.4554	0.1529	2.8167	0.1905	0.2758	0.4975	0.864	0.4675	0.8786
11	AADAC	24.455	35.2528	25.404	17.9134	35.3256	71.6382	22.0789	22.9891	33.0504	28.2588	38.6998	8.5259	14.4697	36.2027
12	AADACL2	23.4643	27.757	22.7782	10.6707	15.9264	20.9236	23.9787	13.2983	20.4217	25.391	25.8656	9.1394	26.6872	25.5244
13	AADACL3	174.8368	94.125	569.6671	205.7836	144.6192	303.7702	390.8507	2.188	812.0157	453.1233	163.1581	423.6692	930.5285	409.5496
14	AADACP1	0	0.0897	0.3623	0.3563	0.1993	0.2	0	0	0.1921	0.2647	0.0597	0.1893	0.0714	0.0615
15	AADAT	4.3083	3.8759	4.5711	4.9101	6.2295	3.7892	5.6352	7.4075	2.4629	2.8587	2.8608	6.6022	2.5542	4.4001
16	AAGAB	30.1732	34.6211	44.2772	29.4888	29.3409	37.0318	31.892	22.3885	30.6468	34.3908	26.699	28.3897	32.6831	28.9904
17	AAK1	4.3885	3.1813	3.9969	5.488	7.4717	2.2428	2.2934	2.5484	1.9751	1.8612	1.9894	3.2594	2.0629	2.7466
18	AAMDC	57.6955	48.2837	29.2412	44.4559	46.0329	40.8307	16.5598	21.1644	12.3302	10.5534	27.2003	26.3361	15.3053	25.0923
19	AAMP	98.7465	101.9495	94.2585	86.2159	100.9133	105.7145	87.5731	80.8099	75.5185	69.417	70.6148	86.6586	83.9103	80.1564



2007年11月05日12:00:32

劳累后左侧腰酸腰痛，口腔溃疡好转，无咽喉疼痛，无左下肢麻木，无肢体浮肿，小便可，舌淡暗，苔薄白腻，脉沉弦。

11.05查尿常规示：PRO3+, BLD1+。

处理：证候改善，效不更方。蛋白量增加，加用ARB。加用丹参活血。

金水宝胶囊[0.33g*63] 3粒/tid/14天/po 2瓶

昆明山海棠片[60片/盒] 2片/tid/14天/po 2盒

厄贝沙坦片(安博维) [0.15g*7s] 0.15g/qd/14天/po 2盒

中药处方：

北芪(黄芪)30g 山萸肉(枣皮/制山萸肉)25g 泽泻(文且/泽且)15g

沙参25g 草决明(决明子)20g 生地(干地/地黄) 15g

茯苓皮(云苓皮)30g 杜仲(盐杜仲)20g 蒲公英20g

霍香(广藿香)15g 淮山(山药)15g 丹皮15g

吐丝子20g 甘草5g 丹参20g

煎服，共7剂

2009年06月24日12时

近1年半在其他肾病专科处就诊。

现无特殊不适，平素易感冒。目前口服强的松25mg qd一年。

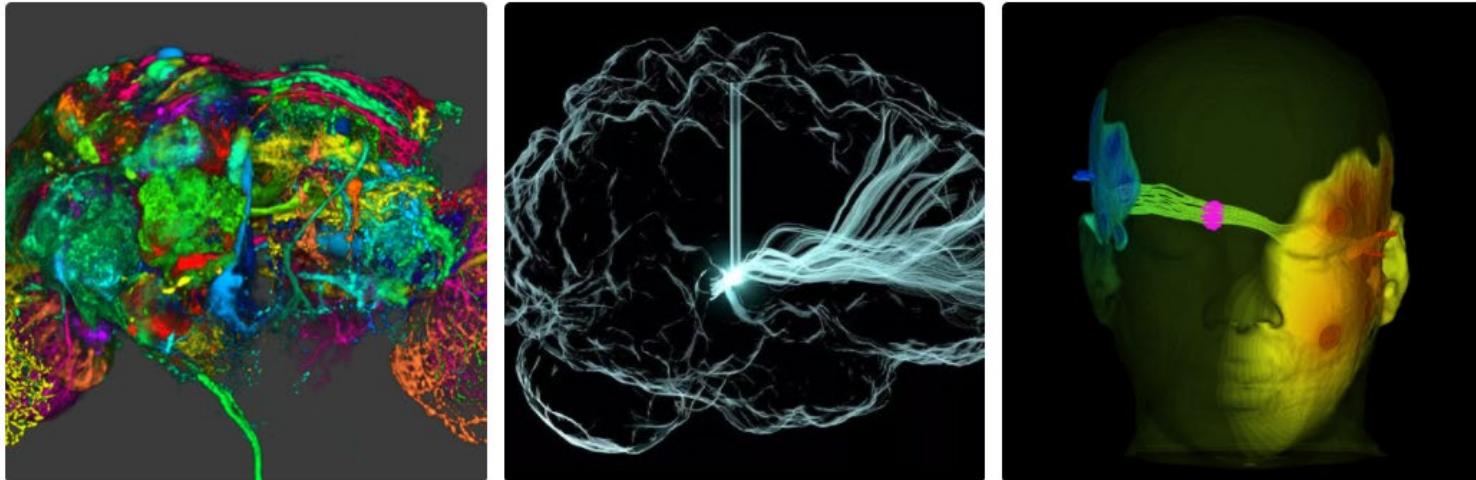
处理：应用丁硼乳膏刷牙预防上呼吸道感染。加用具有类激素样作用的鹿角胶，加用具有抑制免疫炎症作用的苏木、七叶一枝花。

西药处方：

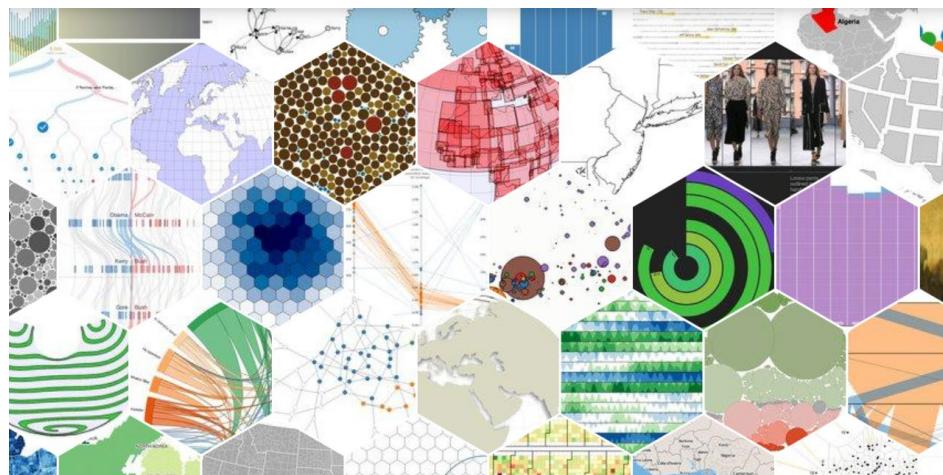


可视化分类

- 科学可视化 Scientific visualization (SciVis)
- 信息可视化 Information visualization (InfoVis)
- 领域内正致力于融合各分支形成完整的学科体系



<https://www.sci.utah.edu/sci-media.html>

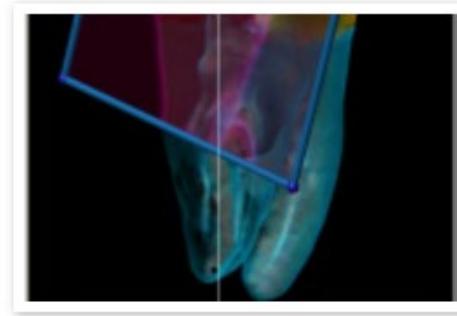
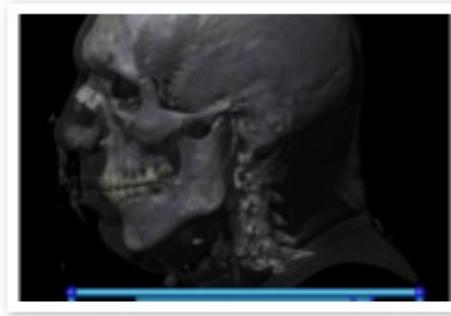
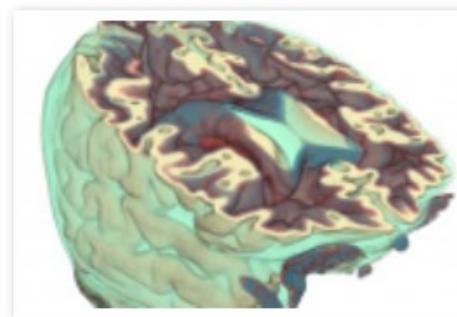
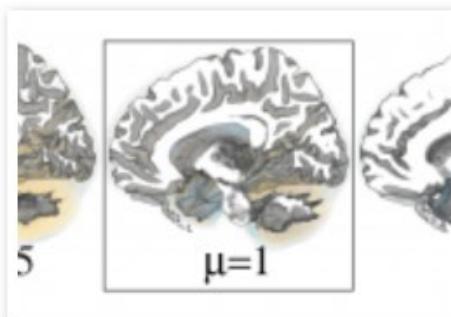
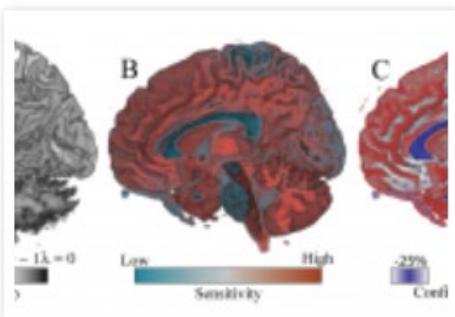
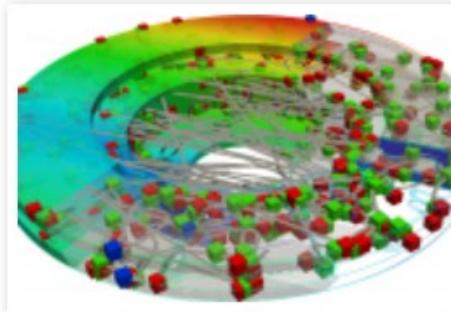
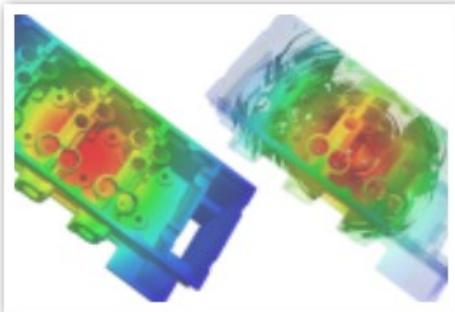


<https://d3js.org/>

©周亮 Liang Zhou

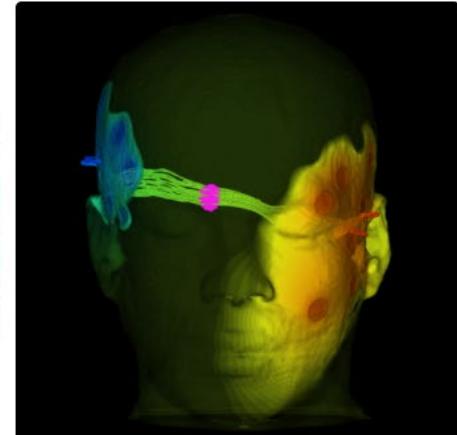
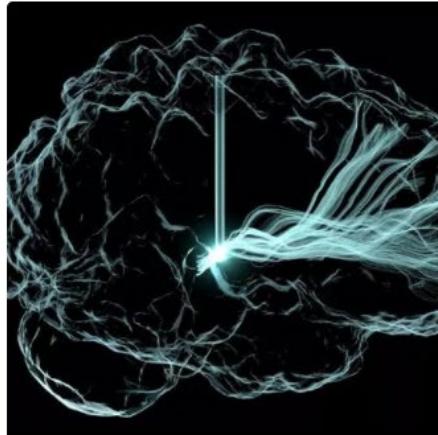
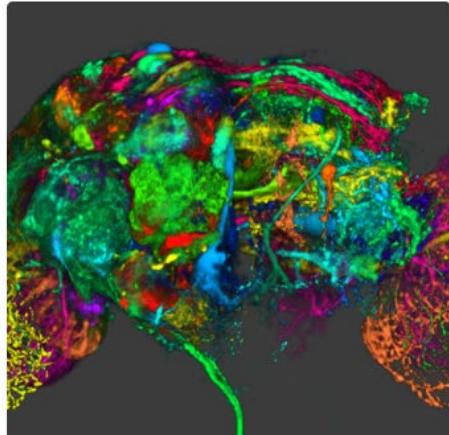
科学可视化

- 研究自带空间信息数据；数据来源来自科学、工程、医学等

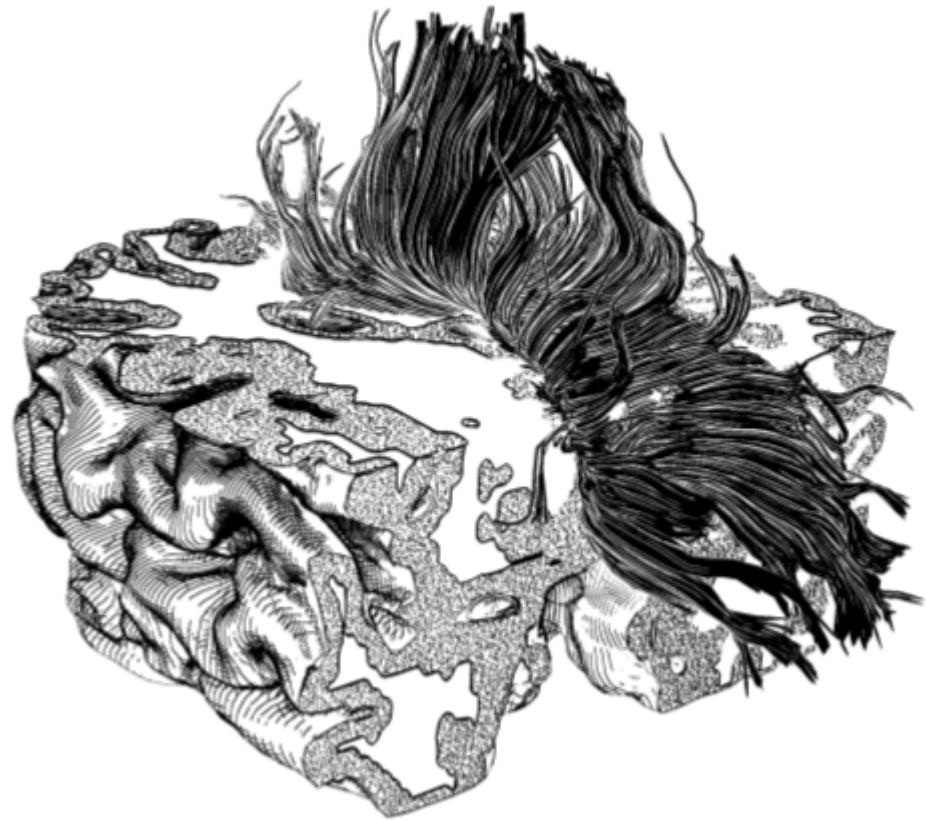
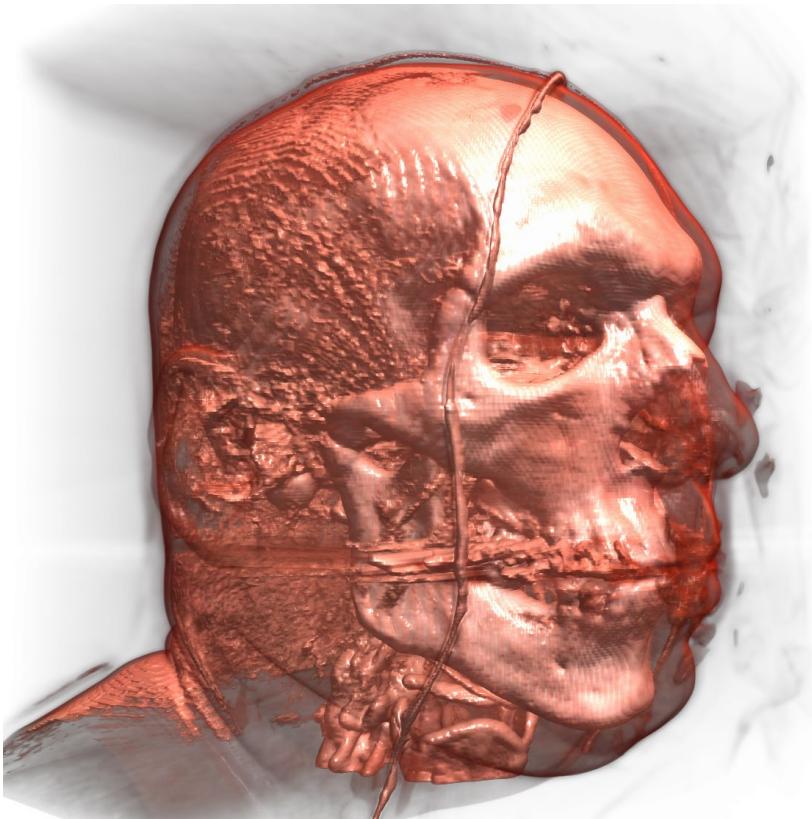


科学可视化

- 多针对假设空间连续的场数据 (field data)
 - 在空间处处有定义, 即, 有数值
- 医学影像 medical imaging
- 生物医疗 biomedical
- 计算流体力学 computational fluid dynamics
- 气象 weather
- 天体物理学 astrophysics
- 地球物理 geophysics
- 能源 energy
-

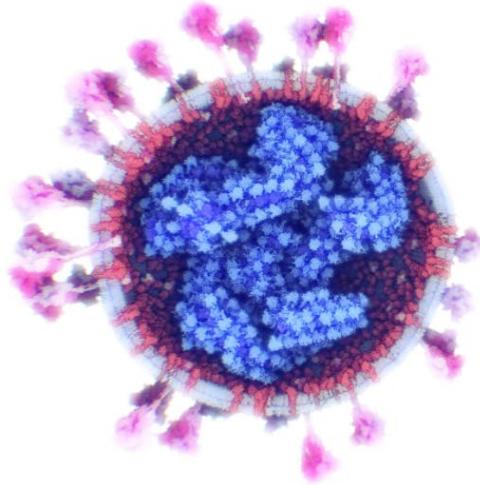


科学可视化——医学影像

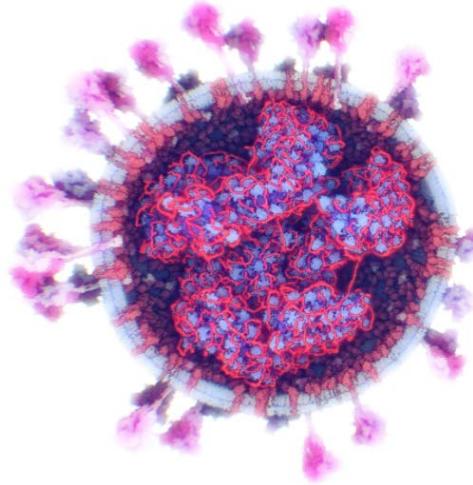


[Pjotr Svetachov, Maarten H. Everts, and
Tobias Isenberg (2010) *DTI in Context:
Illustrating Brain Fiber Tracts In Situ.*]

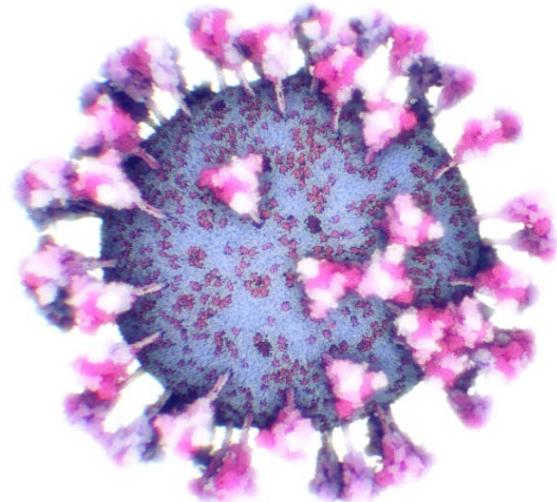
科学可视化——生物医学



Internal nucleoprotein complex

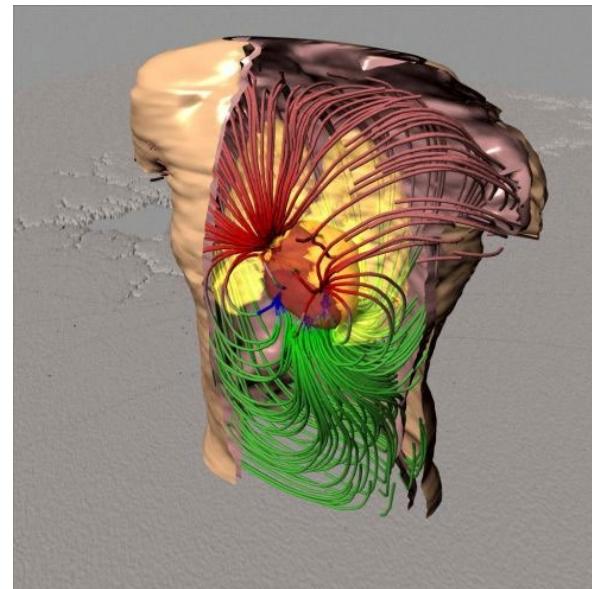


RNA condensed within the nucleoprotein



Outer spike distribution

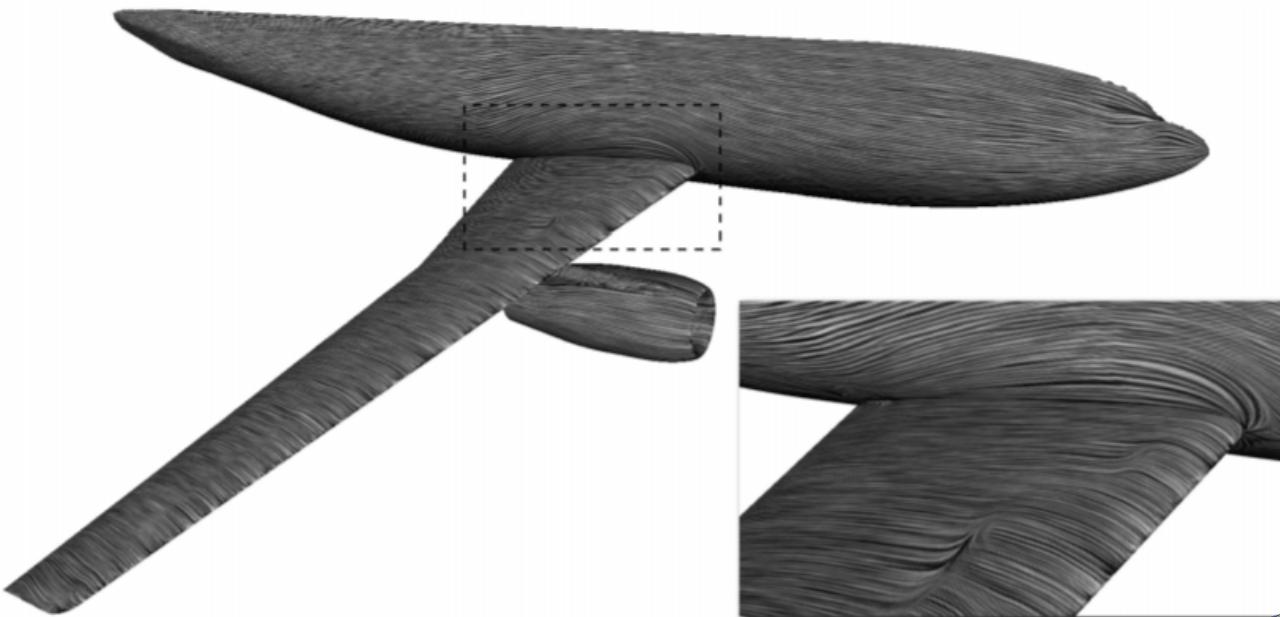
[N. Nguyen *et al.*, "Modeling in the Time of COVID-19: Statistical and Rule-based Mesoscale Models," doi: 10.1109/TVCG.2020.3030415.]



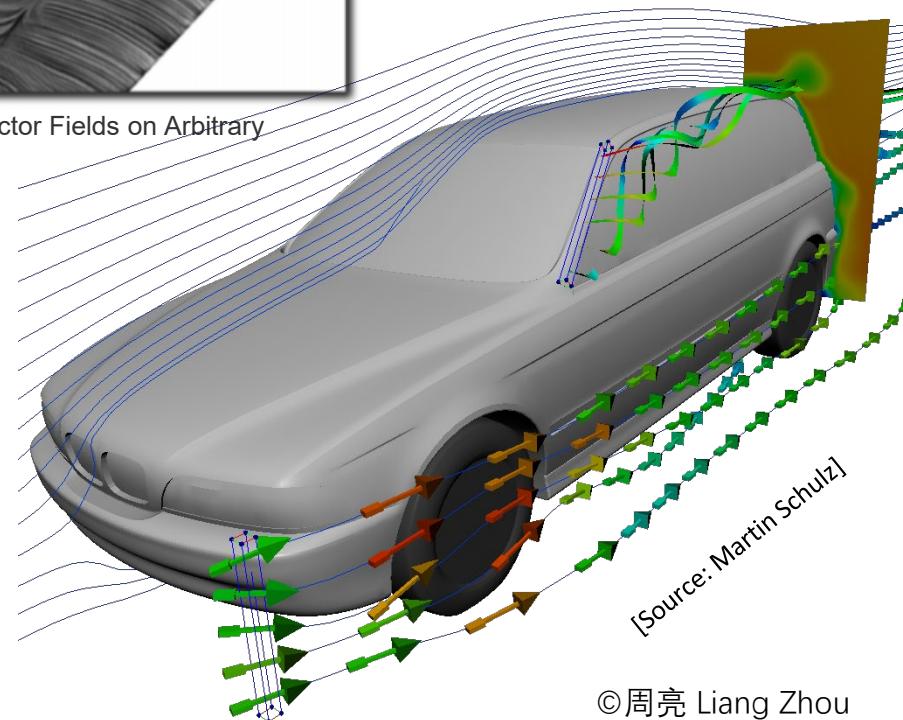
[Bioelectric field modeling with the Utah Torso]

©周亮 Liang Zhou

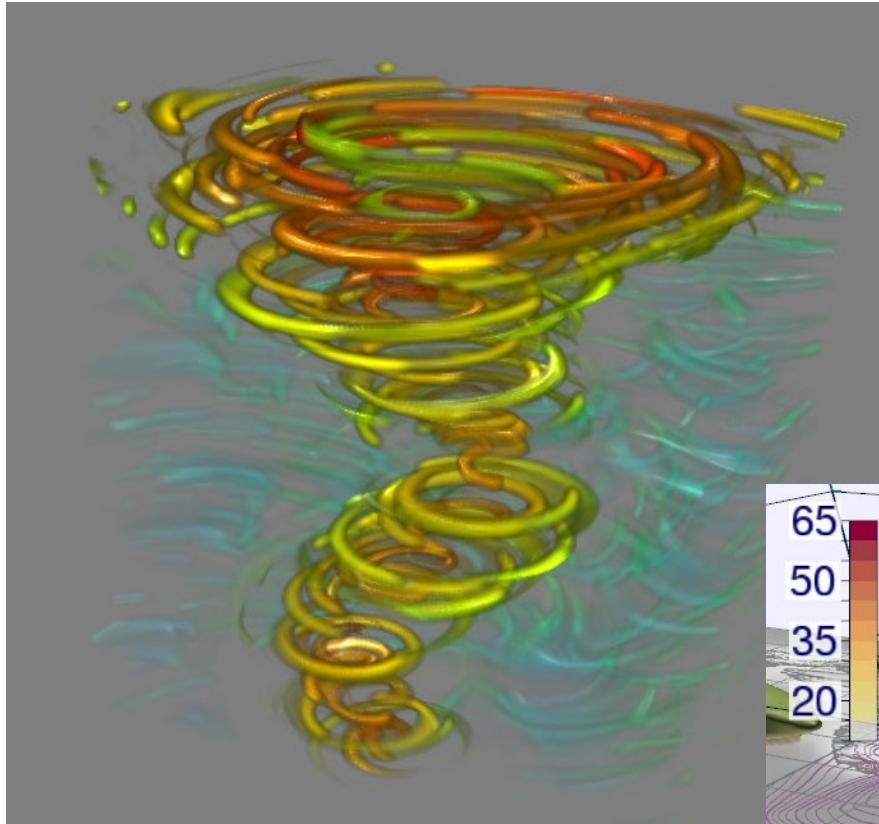
科学可视化——计算流体力学



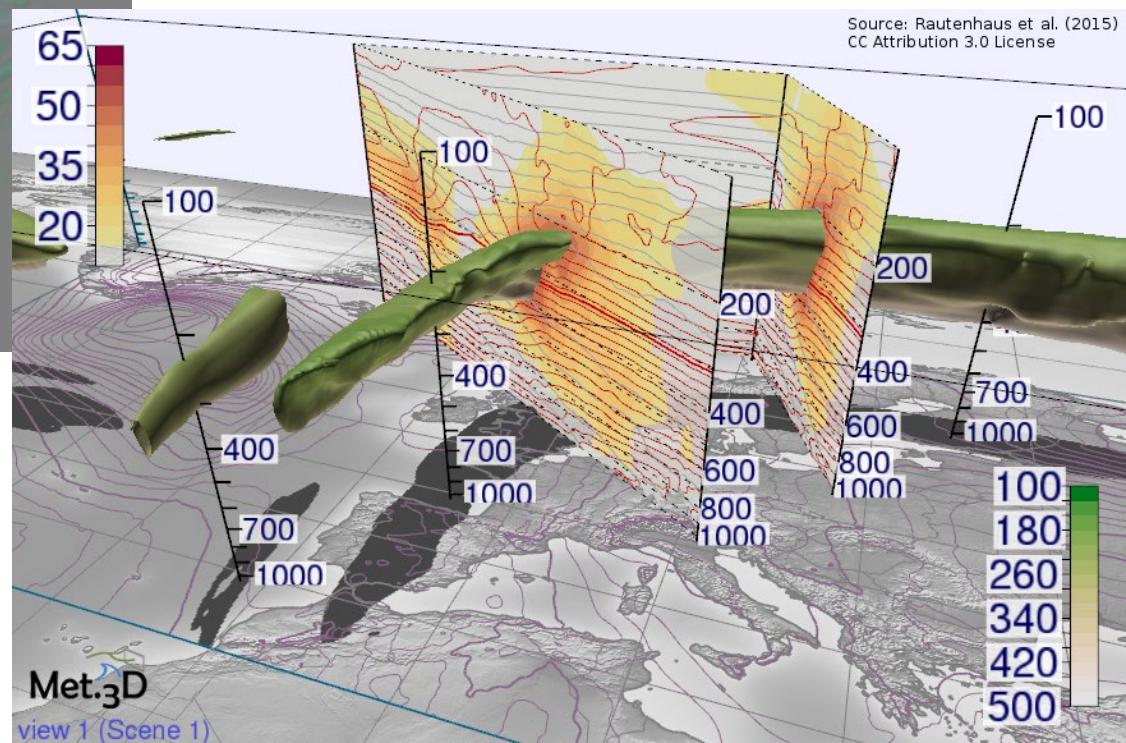
[G. Li, X. Tricoche, D. Weiskopf and C. D. Hansen, "Flow Charts: Visualization of Vector Fields on Arbitrary Surfaces," doi: 10.1109/TVCG.2008.58.]



科学可视化——气象

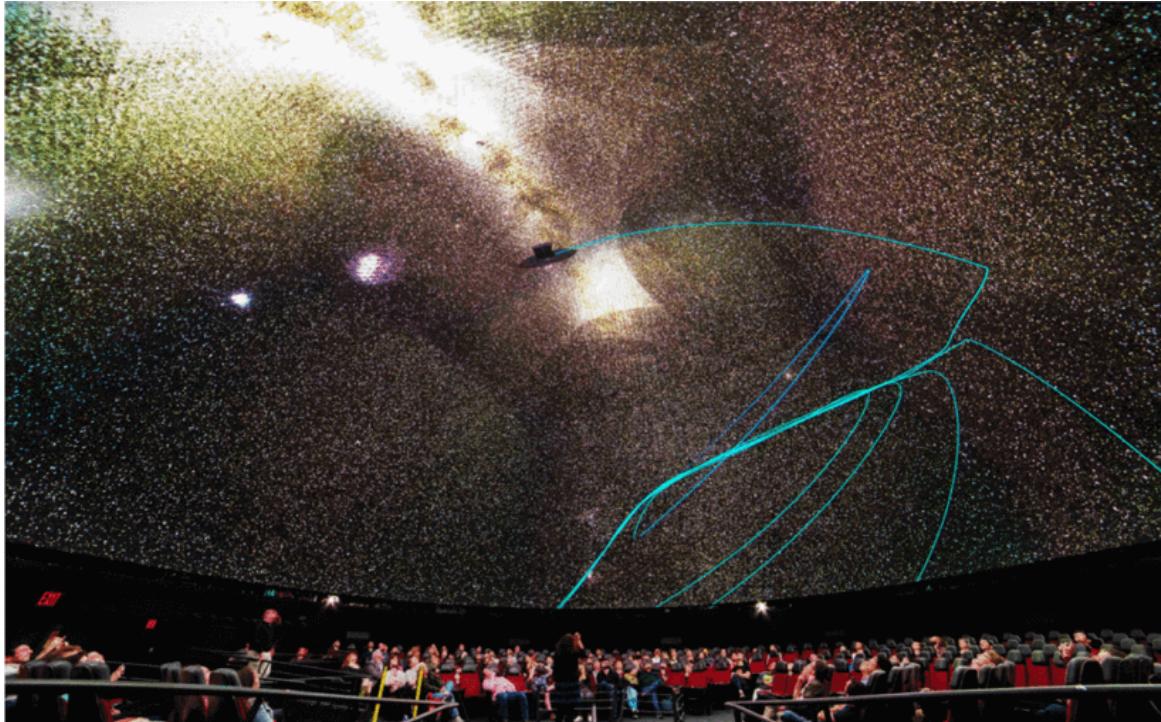


[M. Falk, D. Weiskopf (2008): Output-Sensitive 3D Line Integral Convolution, doi:[10.1109/TVCG.2008.25](https://doi.org/10.1109/TVCG.2008.25)]

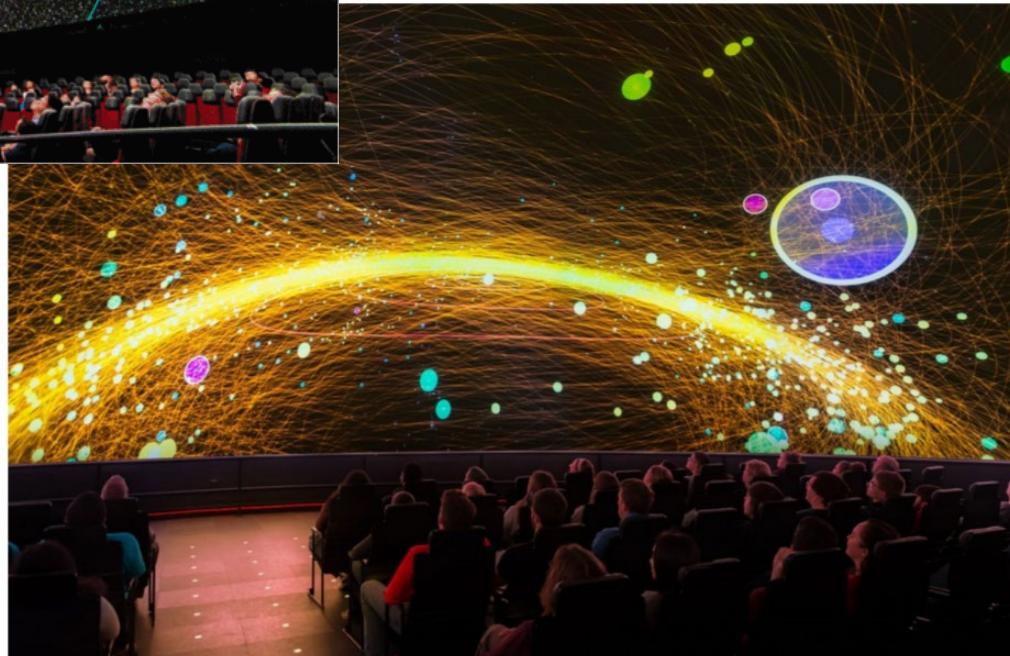


©周亮 Liang Zhou

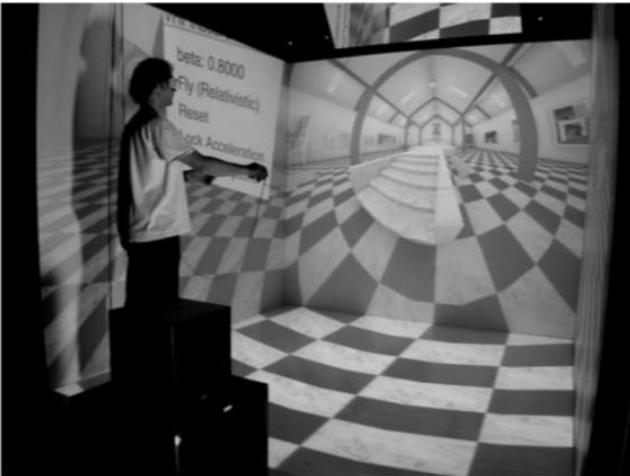
科学可视化——天体物理



[A. Bock et al., "OpenSpace: A System for Astrographics," doi: 10.1109/TVCG.2019.2934259.]



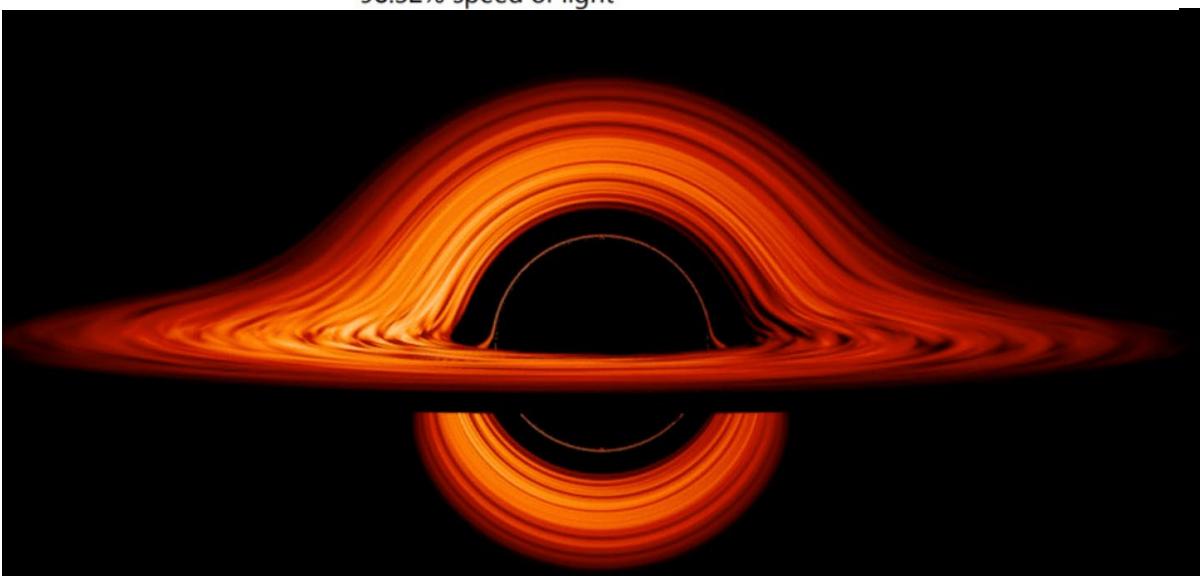
科学可视化——天体物理



80% speed of light



98.32% speed of light



Images: Weiskopf. WSCG 2000

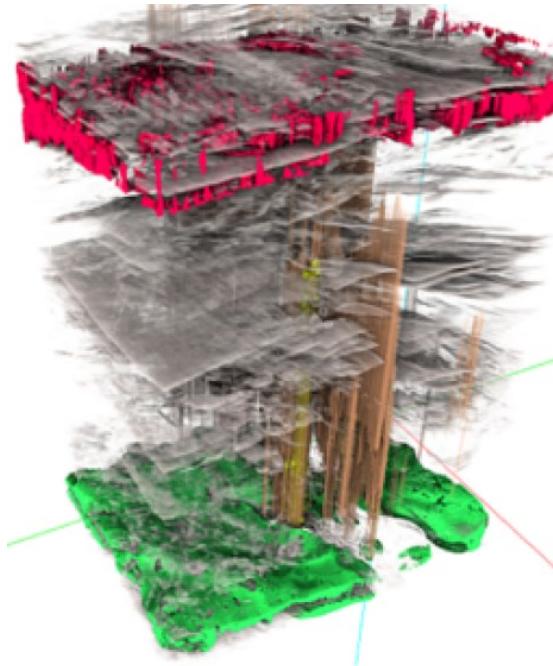
[NASA's Goddard Space Flight Center/Jeremy Schnittman]

周亮 Liang Zhou

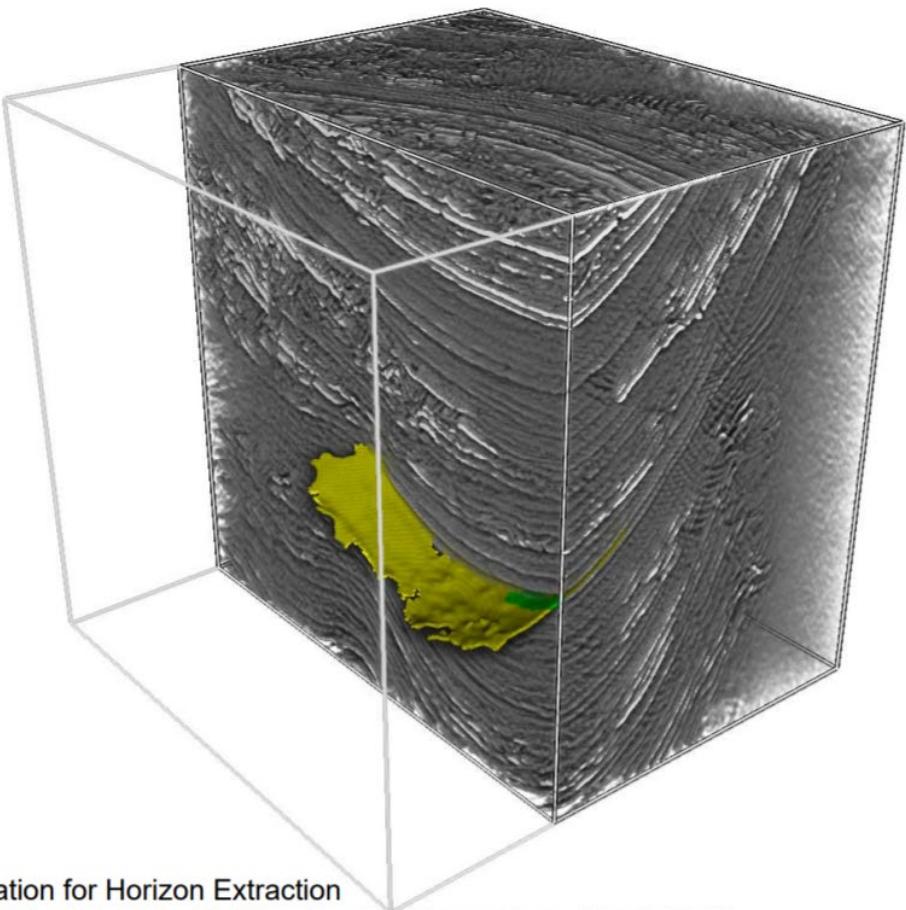


北京大学 健康医疗大数据国家研究院
NATIONAL INSTITUTE OF HEALTH DATA SCIENCE AT PEKING UNIVERSITY

科学可视化——地球物理

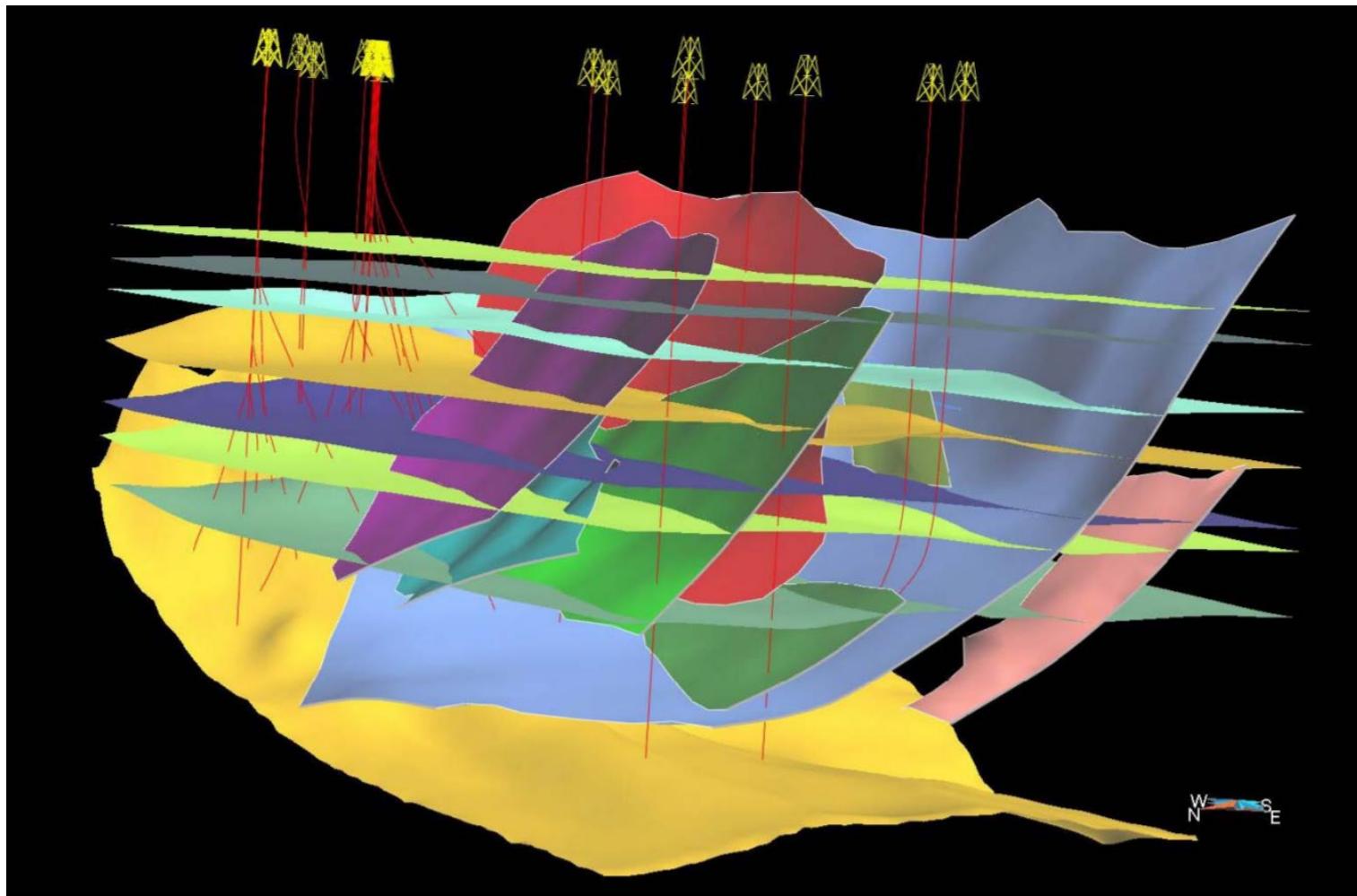


[L. Zhou and C. Hansen, "Transfer function design based on user selected samples for intuitive multivariate volume exploration," doi: 10.1109/PacificVis.2013.6596130]



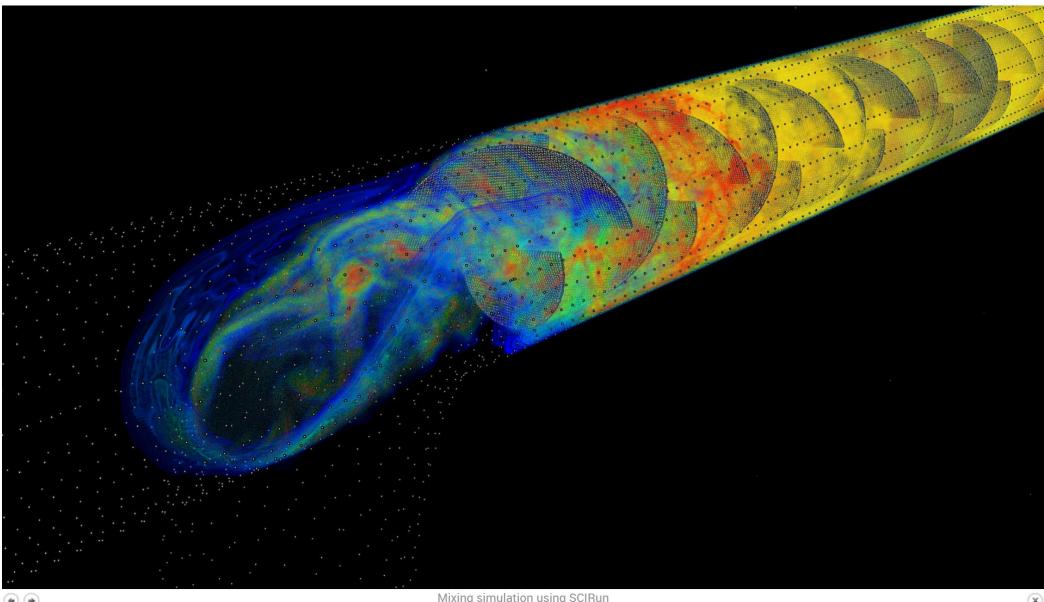
Seismic Volume Visualization for Horizon Extraction
Patel et al. Proceedings of the IEEE Pacific Visualization Symposium. March 2010.

科学可视化——地球物理



VolumeExplorer: Roaming Large Volumes to Couple Visualization and Data Processing for Oil and Gas Exploration. Laurent Castanie et al. Vis 2005

科学可视化——能源

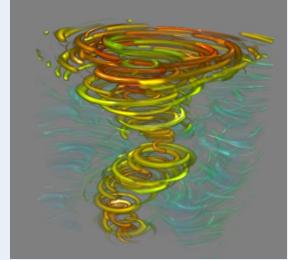
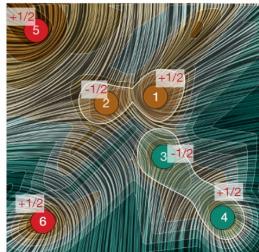
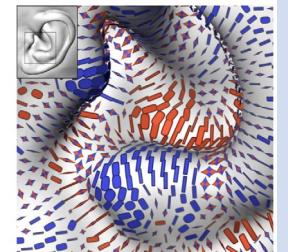


[<https://www.sci.utah.edu/sci-media.html>]

科学可视化

- 根据一个数据变量进行分类
- 假设数据在空间中是连续的——场

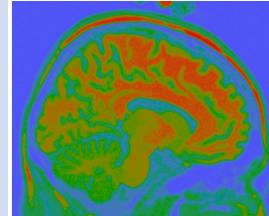
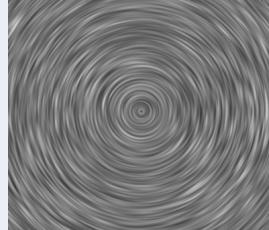
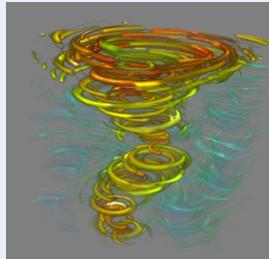
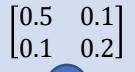
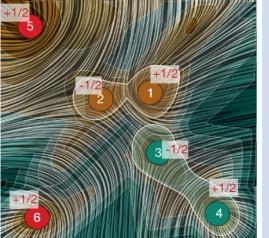
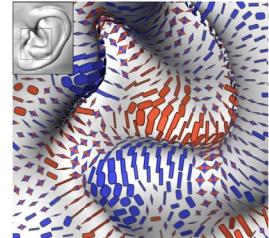
定义域 domain

\mathbb{R}^d	$d = 2$	$d = 3$
\mathbb{R} 标量 scalar 空间上任意一点的值是一个标量		
$\mathbb{R}^m, m = 2, 3$ 向量 vector 空间上任意一点的值是一个向量	 [Source: D. Weiskopf]	 [M. Falk, D. Weiskopf (2008): Output-Sensitive 3D Line Integral Convolution, doi:10.1109/TVCG.2008.25]
$\mathbb{R}^{m \times b}$ 张量 tensor 空间上任意一点的值是一个矩阵	 [J. Jankowiak et al. (2019), Robust Extraction and Simplification of 2D Symmetric Tensor Field Topology, doi:10.1111/cgf.13693]	 [T. Schultz and G. L. Kindermann (2010), "Superquadric Glyphs for Symmetric Second-Order Tensors", doi:10.1109/TVCG.2010.199.]

标量场可视化 Scalar field visualization

- 空间上每个数据点代表一个标量
- 请指出前述例子中哪些是标量场?

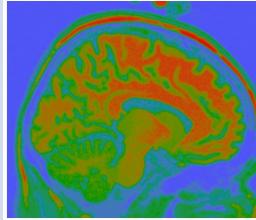
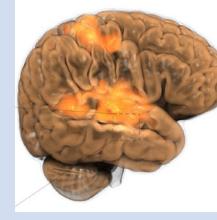
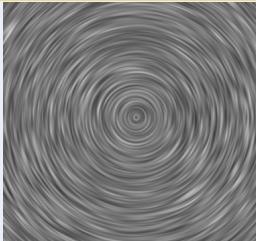
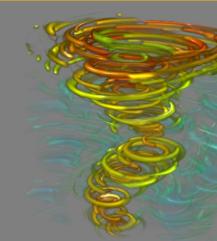
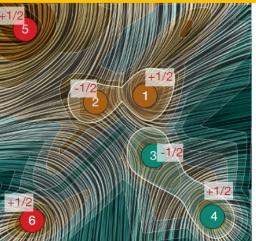
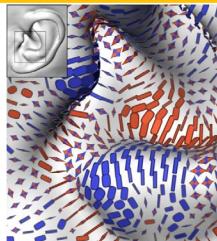
定义域 domain

	\mathbb{R}^d	$d = 2$	$d = 3$
\mathbb{R} 标量 scalar 空间上任意一点的值是一个标量			
$\mathbb{R}^m, m = 2, 3$ 向量 vector 空间上任意一点的值是一个向量			
$\mathbb{R}^{m \times b}$ 张量 tensor 空间上任意一点的值是一个矩阵		 <p>[J. Jankowiak et al. (2019). Robust Extraction and Simplification of 2D Symmetric Tensor Field Topology. doi:10.1111/cgf.13693.]</p>	 <p>[T. Schultz and G. L. Kindermann (2010). "Superquadric Graphics for Symmetric Second-Order Tensors." doi:10.1109/TVCG.2010.199.]</p>

向量场可视化 Vector field visualization

- 空间上每个点的值是一个向量

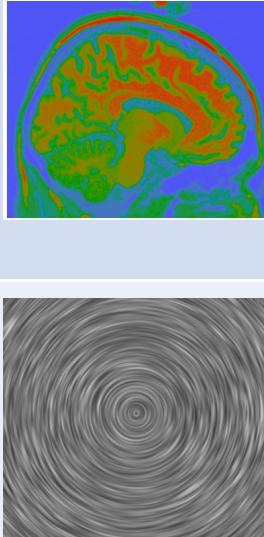
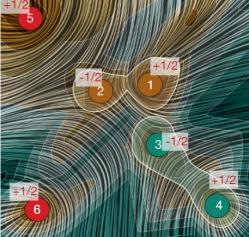
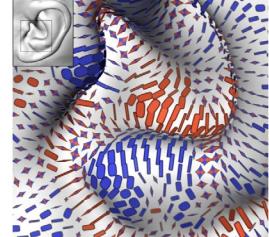
定义域 domain

\mathbb{R}^d	$d = 2$	$d = 3$
\mathbb{R} 标量 scalar 空间上任意一点的值是一个标量 scalar		
$\mathbb{R}^m, m = 2, 3$ 向量 vector 空间上任意一点的值是一个向量	 [Source: D. Weiskopf]	 [M. Falk, D. Weiskopf (2008): Output-Sensitive 3D Line Integral Convolution, doi:10.1109/TVCG.2008.25]
$\mathbb{R}^{m \times b}$ 张量 tensor 空间上任意一点的值是一个矩阵	 [J. Jankowiak et al. (2019): Robust Extraction and Simplification of 2D Symmetric Tensor Field Topology, doi:10.1111/cgf.13693]	 [T. Schultz and G. L. Kindermann (2010), "Superquadratic Graphics for Symmetric Second-Order Tensors," doi:10.1109/TVCG.2010.199.]

张量场可视化 Tensor field visualization

- 空间上每个点的值是一个矩阵 $f: \mathbb{R}^d \rightarrow \mathbb{R}^{m \times b}$

定义域 domain

\mathbb{R}^d	$d = 2$	$d = 3$
\mathbb{R} 标量 scalar 空间上任意一点的值是一个标量		
$\mathbb{R}^m, m = 2, 3$ 向量 vector 空间上任意一点的值是一个向量	 [Source: D. Weiskopf]	 [D. Johnsson and A. Ynnerman (2017). Correlated Photon Mapping for Interactive Global Illumination of Time-varying Volumetric Data. doi:10.1109/TVCG.2016.2598430.]
$\mathbb{R}^{m \times b}$ 张量 tensor 空间上任意一点的值是一个矩阵	 [J. Jankowiak et al. (2019). Robust Extraction and Simplification of 2D Symmetric Tensor Field Topology. doi:10.1111/cgf.13693.]	 [T. Schultz and G. L. Kindermann (2010). "Superquadric Graphics for Symmetric Second-Order Tensors." doi:10.1109/TVCG.2010.199.]

科学可视化研究方向

- 可视化=人+数据+视觉映射+交互

- 绘制远超内存大小的数据
- 并行计算、超级计算
- 视觉上平滑的细节层次过渡

- 调节参数
- 在空间中直接交互
- 混合现实

- 方法有效性用户研究
- 感知实验

- 真实感绘制
- 非真实感绘制
- 视觉感知增强

可视化分类

Low-Pass Filtered Volumetric Shadows

Marco Ament, Filip Sadlo, Carsten Dachsbacher, and Daniel Weiskopf



© 2014

信息可视化 Information visualization

- 可视化不包含固有空间信息的数据——“图表” diagrams,charts, plots
- 缺乏物理世界存在的数据到图形的标准映射
- Information visualization focuses on data sets lacking inherent 2D or 3D semantics and therefore also lacking a standard mapping of the abstract data onto the physical screen space. -- D. Keim

IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 8, NO. 1, JANUARY-MARCH 2002

1

Information Visualization and Visual Data Mining

Daniel A. Keim, Member, IEEE Computer Society

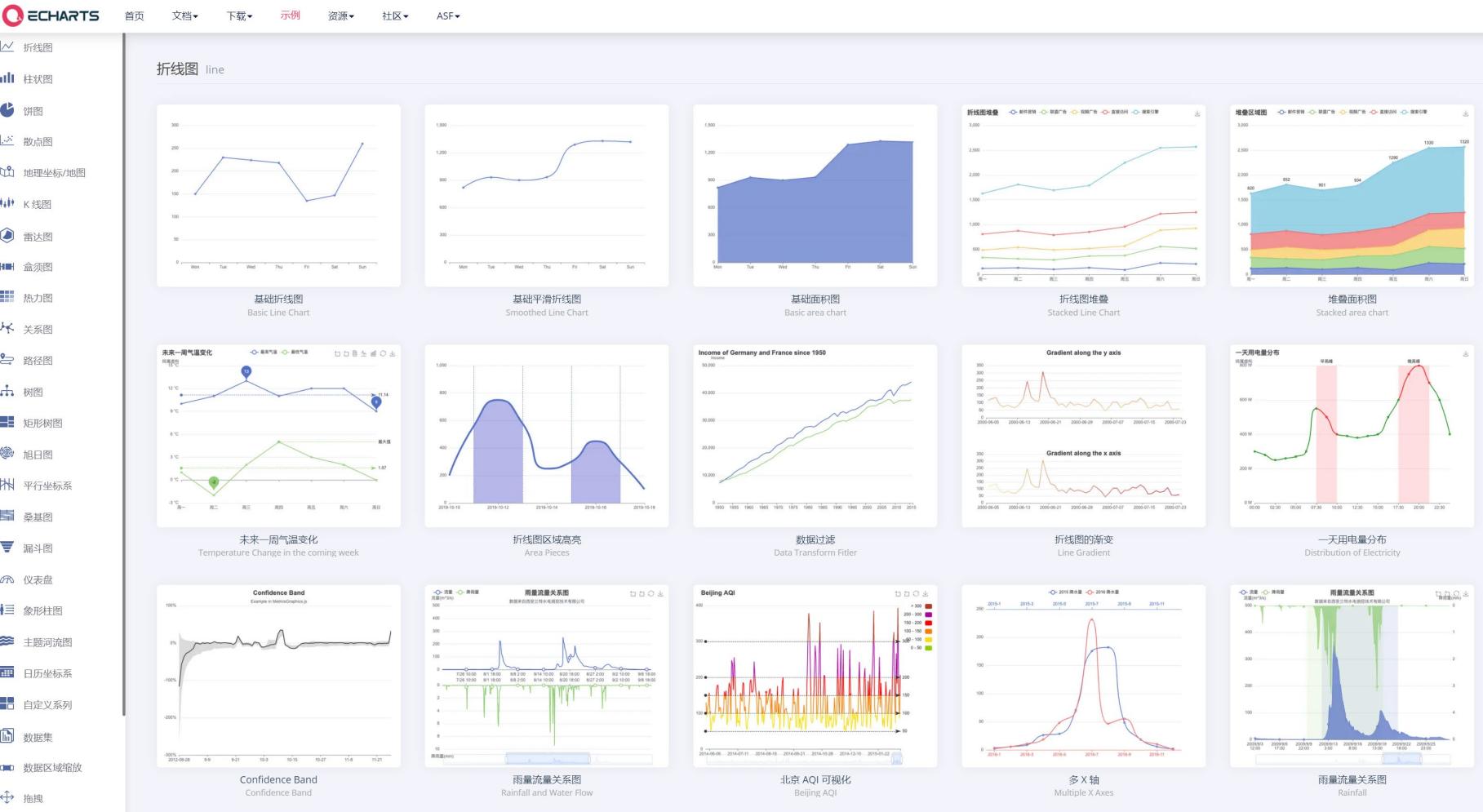
Abstract—Never before in history has data been generated at such high volumes as it is today. Exploring and analyzing the vast volumes of data is becoming increasingly difficult. Information visualization and visual data mining can help to deal with the flood of information. The advantage of visual data exploration is that the user is directly involved in the data mining process. There are a large number of information visualization techniques which have been developed over the last decade to support the exploration of large data sets. In this paper, we propose a classification of information visualization and visual data mining techniques which is based on the *data type to be visualized*, the *visualization technique*, and the *interaction and distortion technique*. We exemplify the classification using a few examples, most of them referring to techniques and systems presented in this special section.

Index Terms—Information visualization, visual data mining, visual data exploration, classification.

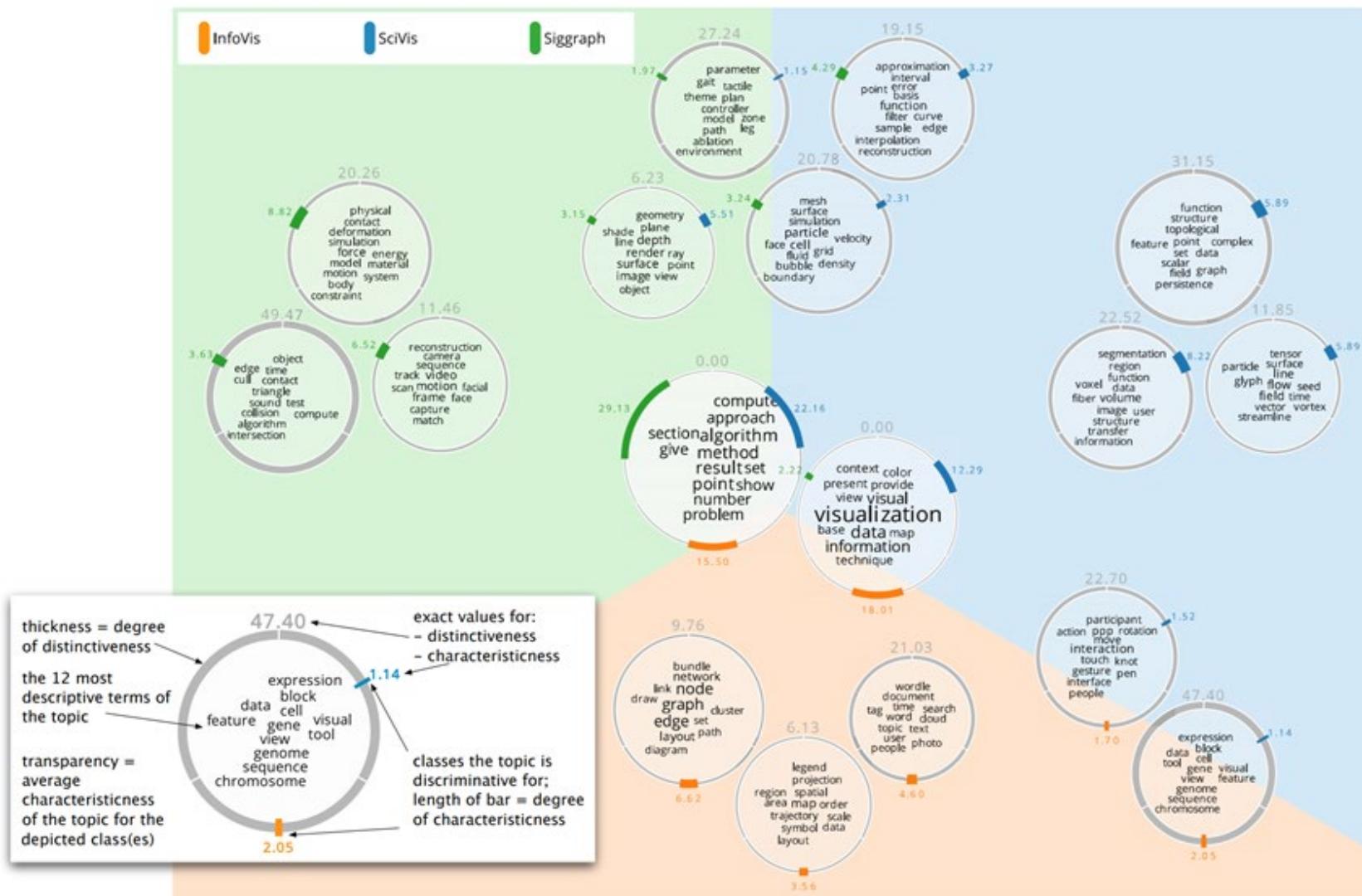
[D. A. Keim, (2002), "Information visualization and visual data mining," doi: 10.1109/2945.981847.]

信息可视化

- 数据来自日常生活、经济、医学、体育、交通 等各方面



与科学可视化的关系



[D. Oelke et al. (2014): Comparative Exploration of Document Collections : a Visual Analytics Approach,
doi: [10.1111/cgf.12376](https://doi.org/10.1111/cgf.12376)]

多样化的数据类型

- 数据集分类
 - 表格数据
 - 树和图
 - 集合
 - 文本
 - 地理数据
 -



A Tour through the Visualization Zoo

A survey of powerful visualization techniques, from the obvious to the obscure

Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky, Stanford University

Thanks to advances in sensing, networking, and data management, our society is producing digital information at an astonishing rate. According to one estimate, in 2010 alone we will generate 1,200 exabytes—60 million times the content of the Library of Congress. Within this deluge of data lies a wealth of valuable information on how we conduct our businesses, governments, and personal lives. To put the information to good use, we must find ways to explore, relate, and communicate the data meaningfully.

The goal of visualization is to aid our understanding of data by leveraging the human visual system's highly tuned ability to see patterns, spot trends, and identify outliers. Well-designed visual representations can replace cognitive calculations with simple perceptual inferences and improve comprehension, memory, and decision making. By making data more accessible and appealing, visual representations may also help engage more diverse audiences in exploration and analysis. The challenge is to create effective and engaging visualizations that are appropriate to the data.

Creating a visualization requires a number of nuanced judgments. One must determine which questions to ask, identify the appropriate data, and select effective *visual encodings* to map data values to graphical features such as position, size, shape, and color. The challenge is that for any given data set the number of visual encodings—and thus the space of possible visualization designs—is

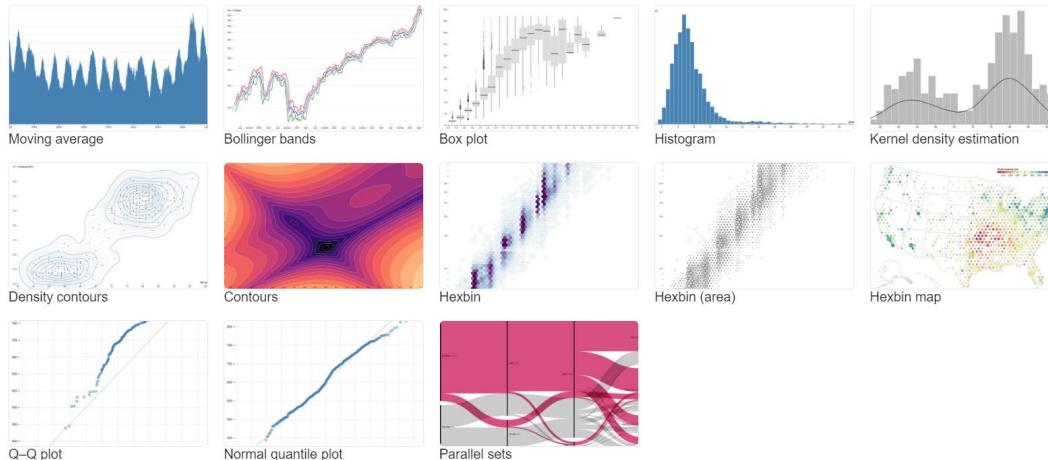
[J. Heer et al. (2010). A Tour through the Visualization Zoo: A survey of powerful visualization techniques, from the obvious to the obscure. DOI: 10.1145/1794514.1805128]

各种信息可视化

▪ 说说你知道的信息可视化？

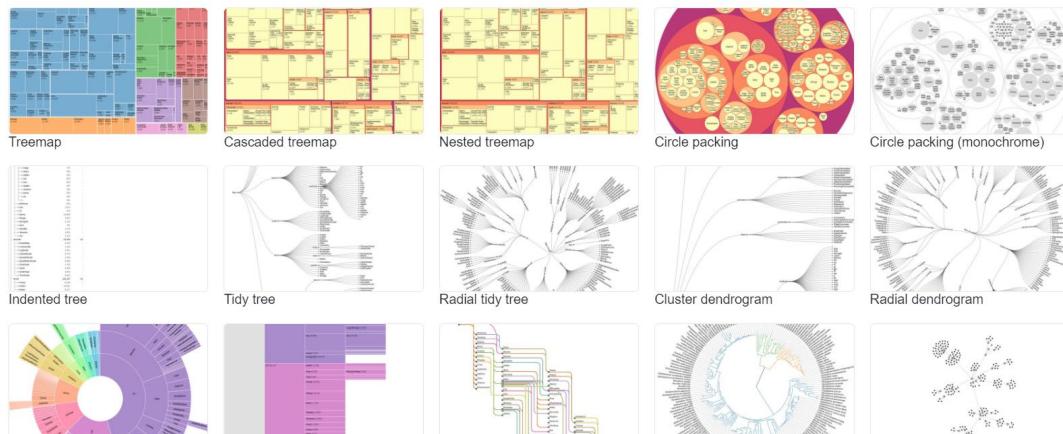
Analysis

D3 is for more than visualization; it includes tools for quantitative analysis, such as [data transformation](#), [random number generation](#), [hexagonal binning](#), and [contours via marching squares](#).

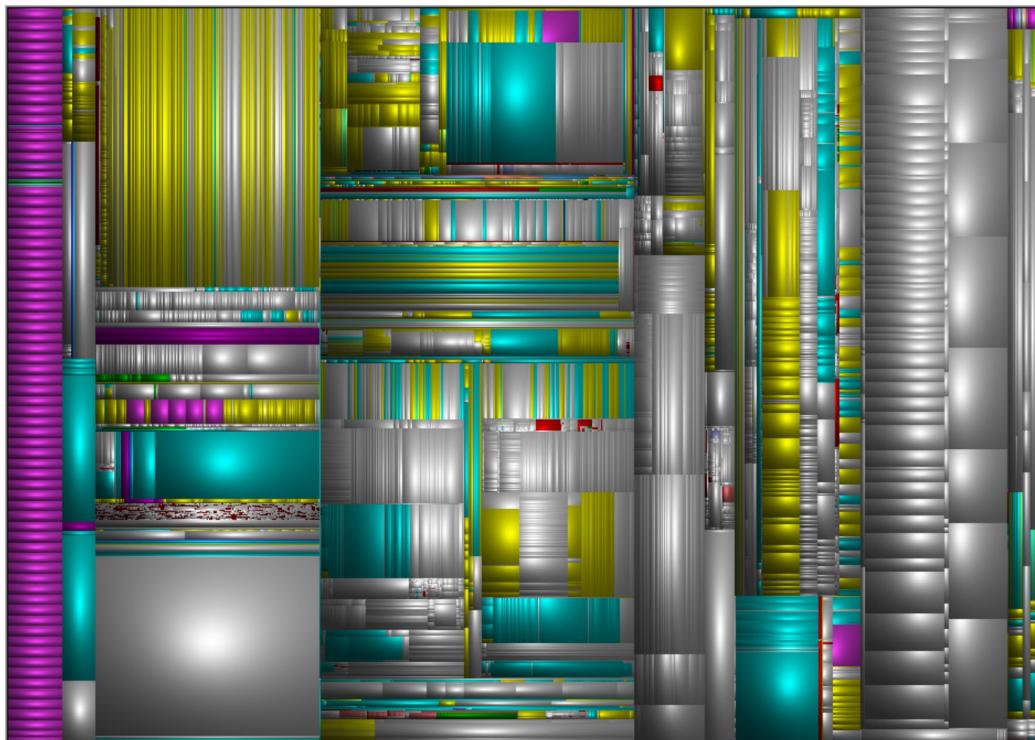


Hierarchies

D3 supports [hierarchical data](#), too, with popular layouts such as [treemaps](#), [tidy trees](#), and [packed circles](#). And you retain complete control over how the data is displayed.



信息可视化——日常生活

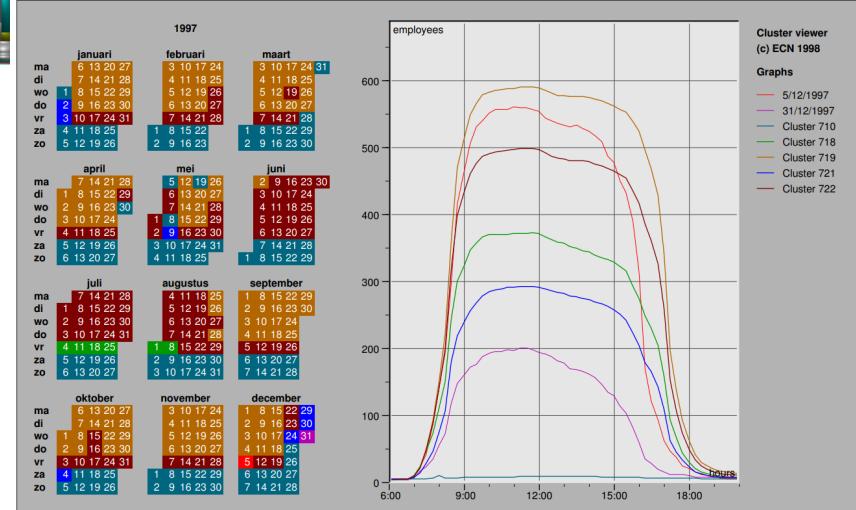


硬盘空间树图

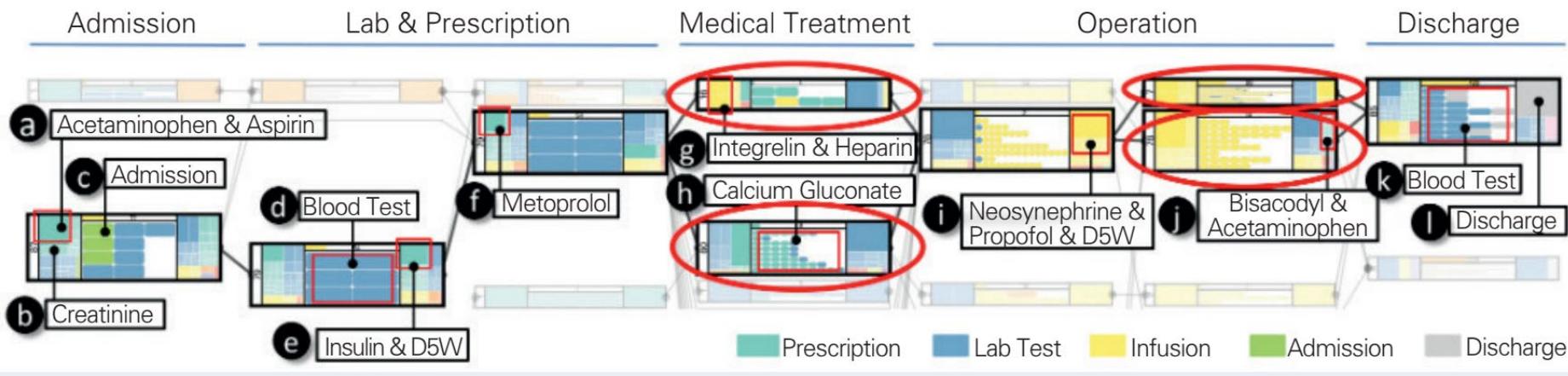
[J.J. van Wijk and H. van de Wetering. Cushion treemaps. In Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99), pages 73–78.]

[J.J. van Wijk and E. van Selow. Cluster and calendar-based visualization of time series data. In Proceedings 1999 IEEE Symposium on Information Visualization (InfoVis'99), pages 4–9.]

公司每日人数规律



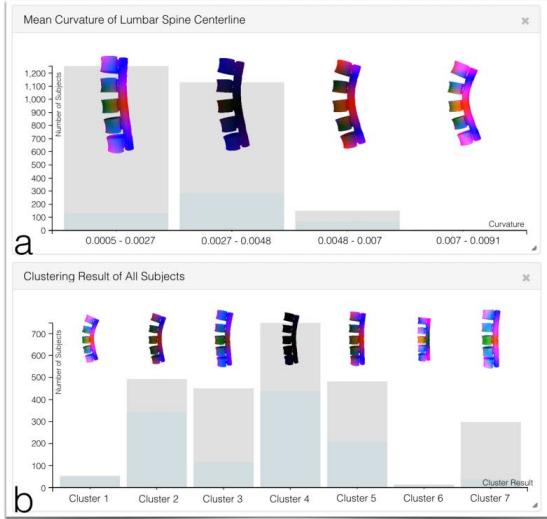
信息可视化——医疗



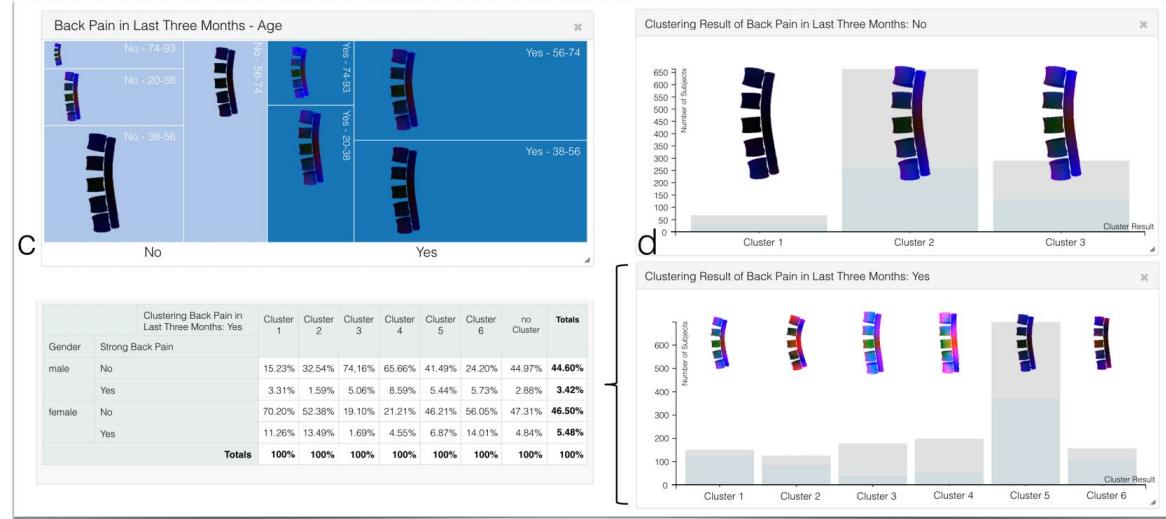
ICU 患者入院序列

[Guo, S., Xu, K., Zhao, R., Gotz, D., Zha, H., & Cao, N. (2018). EventThread: Visual Summarization and Stage Analysis of Event Sequence Data. Doi: 10.1109/TVCG.2017.2745320]

信息可视化——医疗



Case 1: Hypothesis-free Analysis

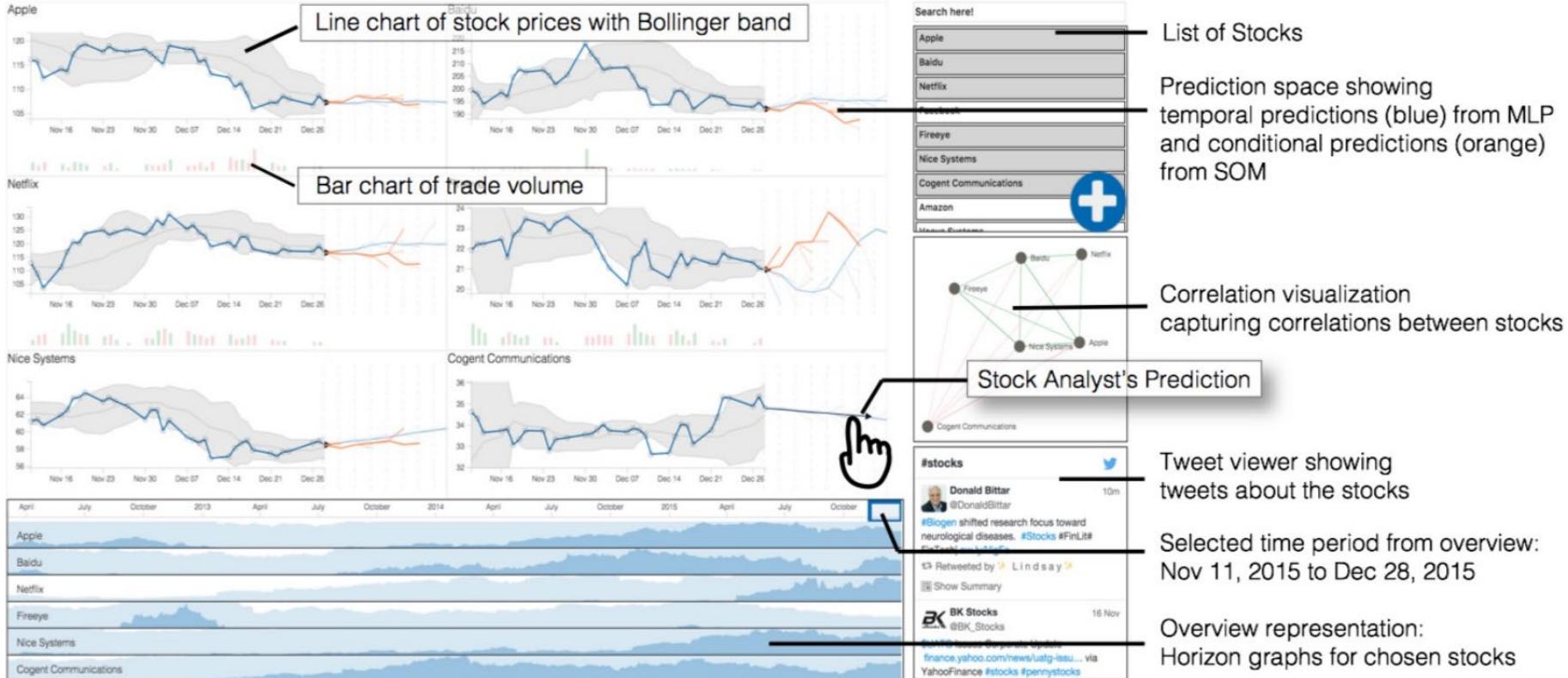


Case 2: Hypothesis-based Analysis

背部疼痛队列研究

[P. Klemm, S. Oeltze-Jafra, K. Lawonn, K. Hegenscheid, H. Völzke and B. Preim, "Interactive Visual Analysis of Image-Centric Cohort Study Data," , doi: 10.1109/TVCG.2014.2346591.]

信息可视化——经济

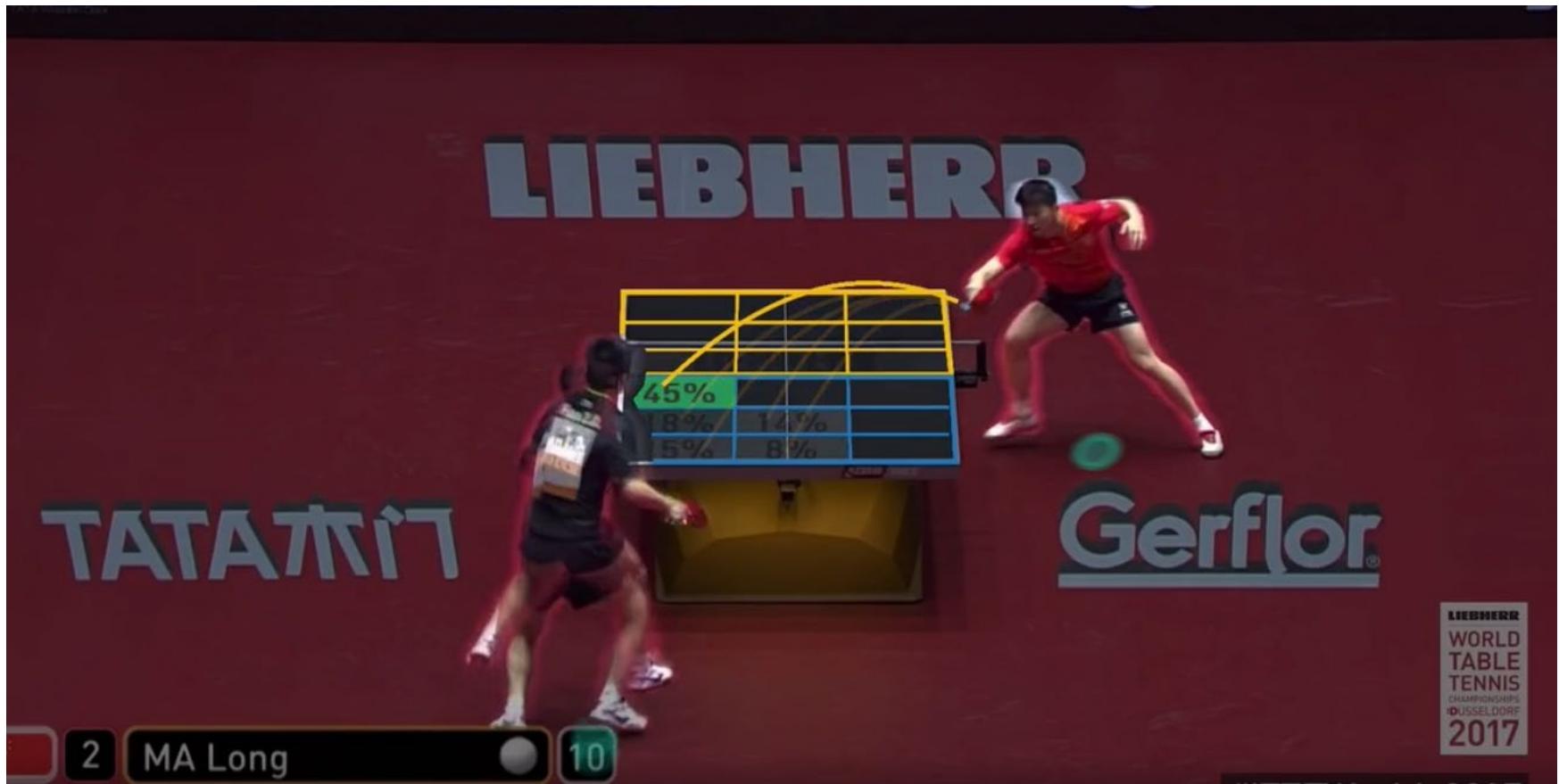


股价分析与预测

[Badam, S. K., Zhao, J., Sen, S., Elmquist, N., & Ebert, D. (2016). Timefork: Interactive prediction of time series. doi: 10.1145/2858036.2858150.]

信息可视化——体育

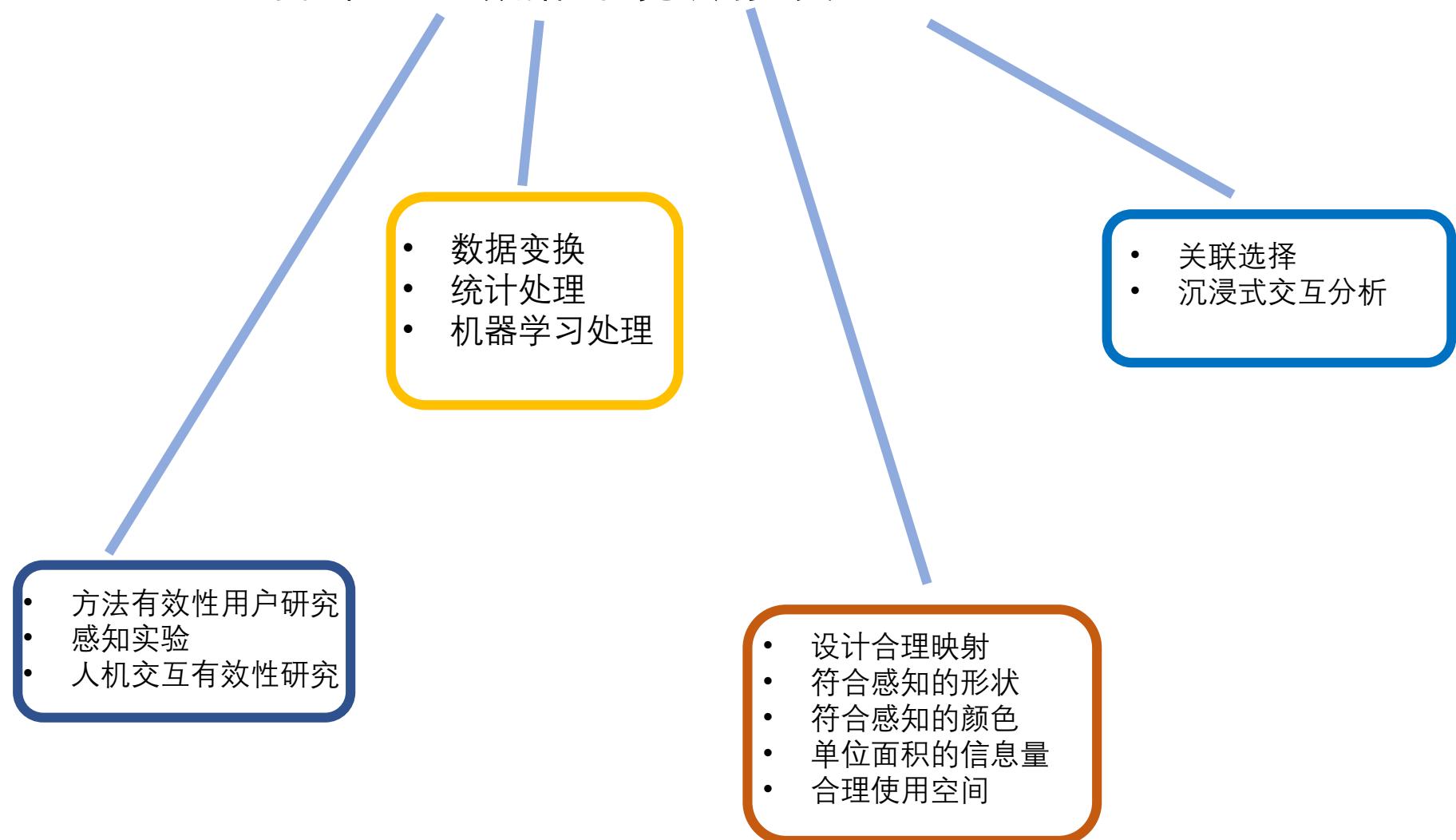
Augmenting Sports Videos with VisCommentato



<https://viscommentator.github.io/#examples>

信息可视化研究方向

- 可视化=人+数据+视觉映射+交互

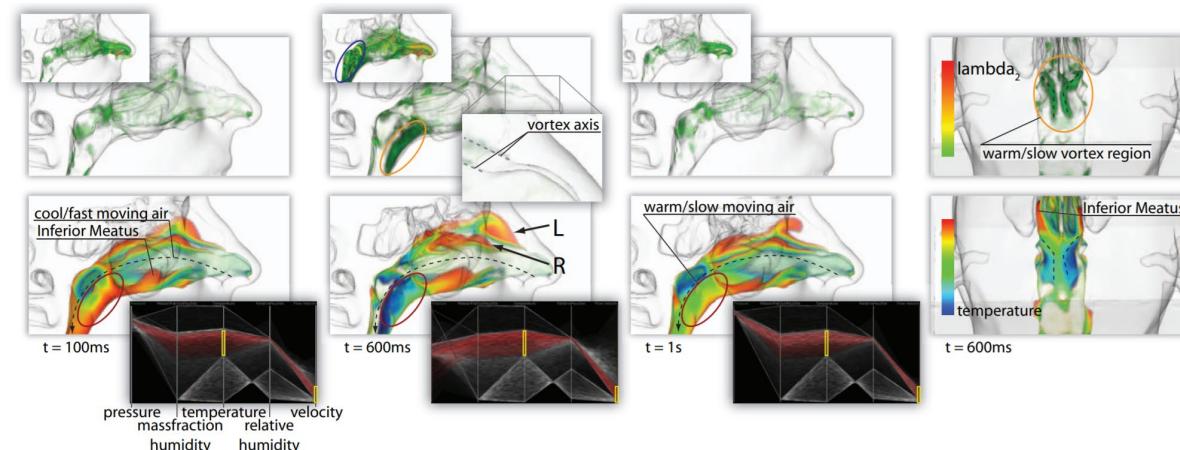


可视分析 Visual Analytics

- Visual analytics is the science of analytical reasoning supported by interactive visual interfaces.

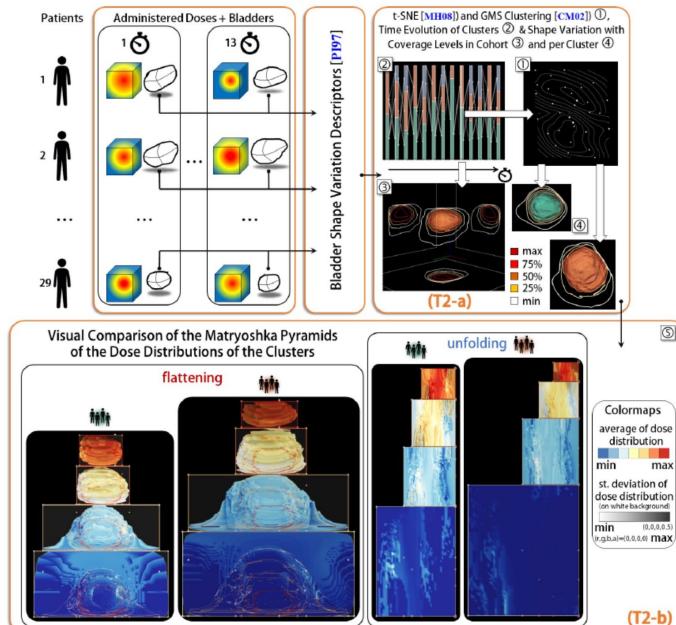
——[Keim D.A. et al. (2009) Visual Analytics. DOI: 10.1007/978-0-387-39940-9_1122]

- 可视化研究对一类数据进行绘制、展示、交互的方法
- 可视分析用集成多种可视化方法的交互可视界面解决具体问题



鼻腔气流分析

[S. Zachow et al., "Visual Exploration of Nasal Airflow," doi: 10.1109/TVCG.2009.198.]



放疗导致膀胱中毒分析

[R. Raidou et al. "Bladder Runner: Visual Analytics for the Exploration of RT-Induced Bladder Toxicity in a Cohort Study,". doi:10.1111/cgf.13413]

©周亮 Liang Zhou

可视分析

- 有机结合各种可视化，交互地进行数据分析
- User Interaction 用户交互
- Brushing and linking 关联更新
- Linked views 关联视图
- Focus + context 聚焦加上下文技术

Jeffrey Heer and Ben Shneiderman. 2012.
Interactive dynamics for visual analysis.
Commun. ACM 55, 4 (April 2012), 45–54.
DOI:10.1145/2133806.2133821

DOI:10.1145/2133806.2133821

 Article development led by **queue.acm.org**
queue.acm.org

**A taxonomy of tools that support the fluent
and flexible use of visualizations.**

BY JEFFREY HEER AND BEN SHNEIDERMAN

Interactive Dynamics for Visual Analysis

THE INCREASING SCALE and availability of digital data provides an extraordinary resource for informing public policy, scientific discovery, business strategy, and even our personal lives. To get the most out of such data, however, users must be able to make sense of it: To pursue questions, uncover patterns of interest, and

可视化各个子领域的融合

Restructuring

The VIS Restructuring Committee (2016-2018)

The VIS Restructuring Committee was tasked in 2016 by the VEC (VIS Executive Committee) with considering alternative structures, such as a more unified conference, that may better enhance vibrancy and growth. It has generated the following documents:

- [VIS Restructuring Recommendations, Fall 2018](#)
- [VIS Restructuring Workshop Executive Summary, Fall 2018](#)
- [VIS Restructuring Workshops Summary, Fall 2018](#)
- [VIS Restructuring Feedback, January 2018](#)
- [VIS Restructuring Report \(“Phoenix Report”\), Fall 2017](#)

Committee Members: Hanspeter Pfister (chair), Hans Hagen, Daniel Keim, Tamara Munzner, Stephen North

The VIS Restructuring Committee has concluded its work and has spawned new committees with a more specific mandate in 2019.

The reVISe Committee (2019-ongoing)

The reVISe Committee was struck in February 2019 by the VEC (VIS Executive Committee), with the mandate of creating a concrete proposal for re-organizing the academic conferences (VAST, InfoVis, SciVis) under a single umbrella, including paper reviewing and governance. It meets weekly, and posts intermediate work products regularly including summarized meeting minutes and key findings. To contact the committee with feedback on any aspect of its work, please write to revise@ieeveis.org.

推荐阅读

- The Value of Visualization. Jarke van Wijk. Proceedings of the IEEE Visualization Conference, pp. 79-86, 2005.

https://hci.stanford.edu/courses/cs448b/papers/vanwijk_vov.pdf

The Value of Visualization

Jarke J. van Wijk*
Dept. Mathematics and Computer Science
Technische Universiteit Eindhoven

ABSTRACT

The field of Visualization is getting mature. Many problems have been solved, and new directions are sought for. In order to make good choices, an understanding of the purpose and meaning of visualization is needed. Especially, it would be nice if we could assess what a good visualization is. In this paper an attempt is made to determine the value of visualization. A technological viewpoint is adopted, where the value of visualization is measured based on effectiveness and efficiency. An economic model of visualization is presented, and benefits and costs are established. Next, consequences for and limitations of visualization are discussed (including the use of alternative methods, high initial costs, subjectiveness, and the role of interaction), as well as examples of the use of the model for the judgement of existing classes of methods and understanding why they are or are not used in practice. Furthermore, two alternative views on visualization are presented and discussed: viewing visualization as an art or as a scientific discipline. Implications and future directions are identified.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces; I.3.6 [Computer Graphics]: Methodology and Techniques I.3.8 [Computer Graphics]: Applications

Keywords: Visualization, evaluation

1 INTRODUCTION

In this paper I want to give a contribution to the discussion on the status and possible directions of our field. Rather than to pinpoint specific topics and activities, my aim is to detect overall patterns, and to find a way to understand and qualify visualization in general. This is an ambitious and vague plan, although the basic ground for this is highly practical.

I have to make decisions on visualization in many roles. As a researcher, decisions have to be made ranging from which area to spend time on to which particular solution to implement; as a supervisor, guidance to students must be provided; as a reviewer, new results and proposals for new research must be judged, and opinions are expected if they are worth publishing or funding; as advisor in a start-up company, novel and profitable directions must be spotted. All these cases imply judgement of the value of visualization in varying senses.

How to assess the value of visualization? Visualization itself is an ambiguous term. It can refer to the research discipline, to a technology, to a specific technique, or to the visual result. If visualization is considered as a technology, i.e., as a collection of methods, techniques, and tools developed and applied to satisfy a need, then standard measures apply: Visualization has to be *effective* and *efficient*. In other words, visualization should do what it is supposed to do, and has to do this using a minimal amount of resources. One immediate and obvious implication is that we cannot judge visualization on its own, but have to take into account the context in which it is used.

In section 2 a short overview is given of the background of the