可視化の世界

計算科学が拓く世界

2017年4月26日（水）

学術情報メディアセンター
小山田耕二

内容

- 可視化について
- 可視化事例
- 因果推論技術
- 課題

本日のテーマと目的

- 可視化は、計算機や計測装置等から生成される膨大な数値データから気付きを得るための基盤技術として重要になっている。本講義では、計算科学と密接な関係にある可視化技術の基礎と応用について説明する。

映画にみる人工知能（AI）のもたらす終末期

- ターミネータ
- バイオハザード

Hot Robot Wants To Destroy Humans

- Robotics is finally reaching the mainstream and androids - humanlike robots - are everywhere at SWSX

機械に奪われそうな仕事

http://blog.btrax.com/jp/2016/01/18/ai-skills/

- プログラマー: **48.1%**
- ソフトウェアエンジニア: **4.2%**
- 家政婦: **68.8%**
- ウェイター/ウェイトレス: **93.7%**
- バーテン: **76.8%**
- 調理師: **96.3%**
- シェフ: **10.1%**
- 経理: **97.6%**
- 経理部長: **6.9%**

知性増幅(IA)と可視化

- AIは、ビッグデータを使って機械に学習させるのに役立つ
- ビッグデータを使ってヒトに学習させるのに役立つのは？

計算科学が拓く世界

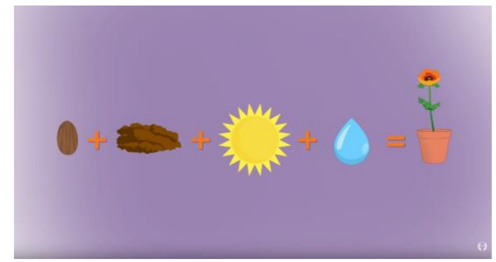
科学的方法習得の低年齢化

https://www.youtube.com/watch?v=KIFz_-KzURY

Scientific Method Song Video

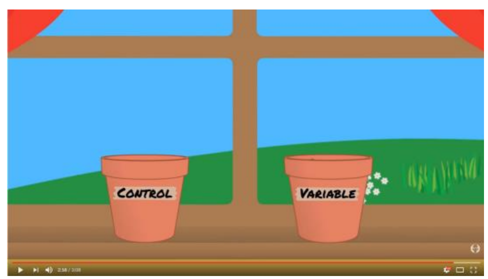
Make an observation

Ask a question

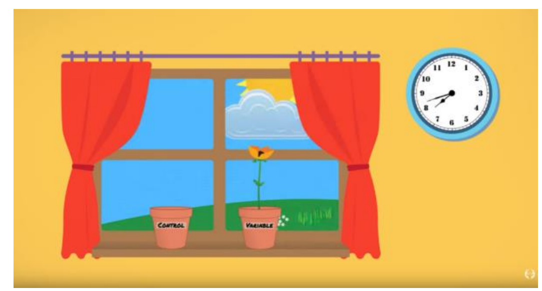


Form a hypothesis

Make a prediction



Do a test or experimentation



Analyze data and draw a conclusion

Dance

可視化について

ビッグデータ時代の科学を支える素養

Source: IDC's Digital Universe Study, sponsored by EMC, June 2011

1.8 zeta bytes in 2011

データと情報

データ

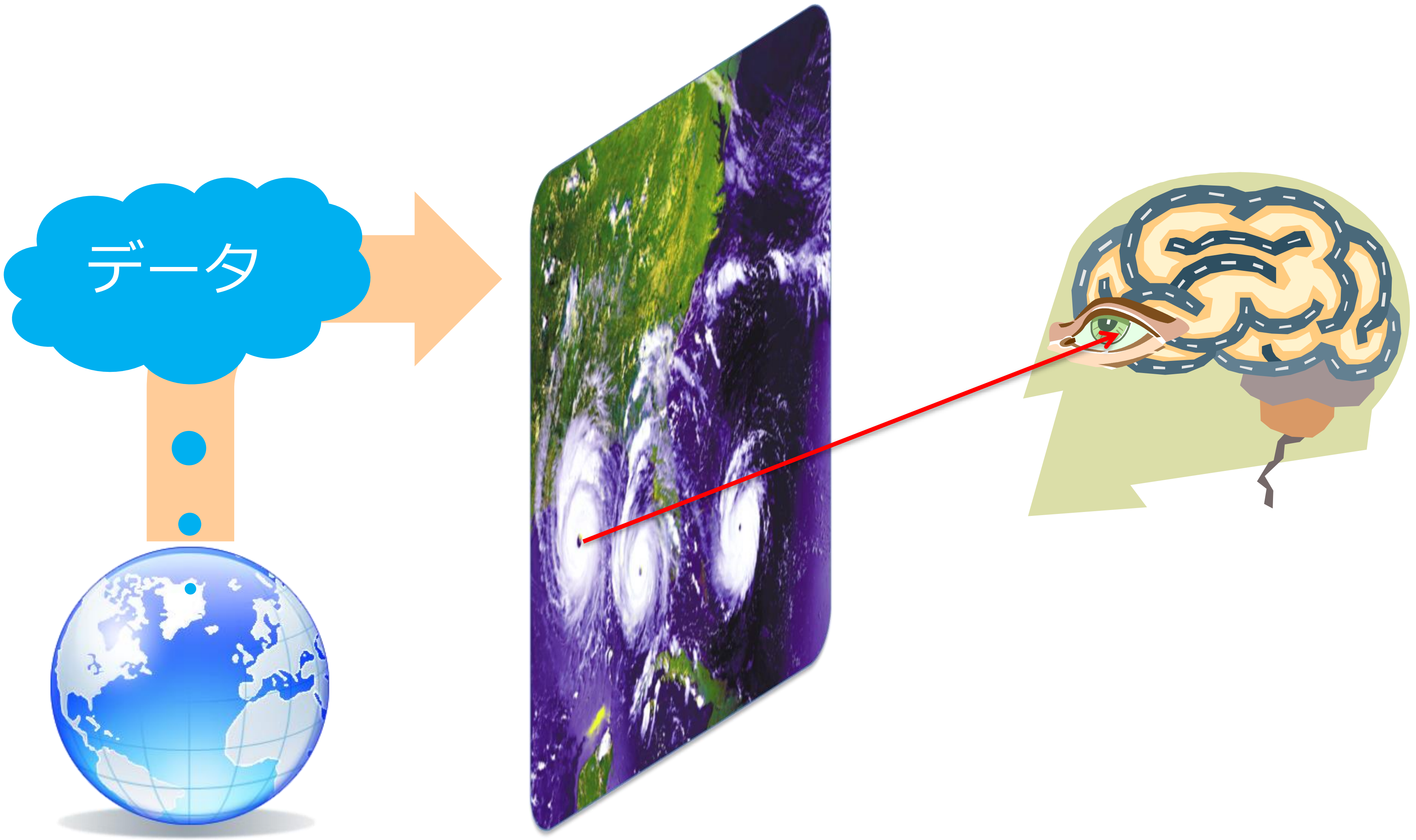
何かを符号で表現したもの

可視化

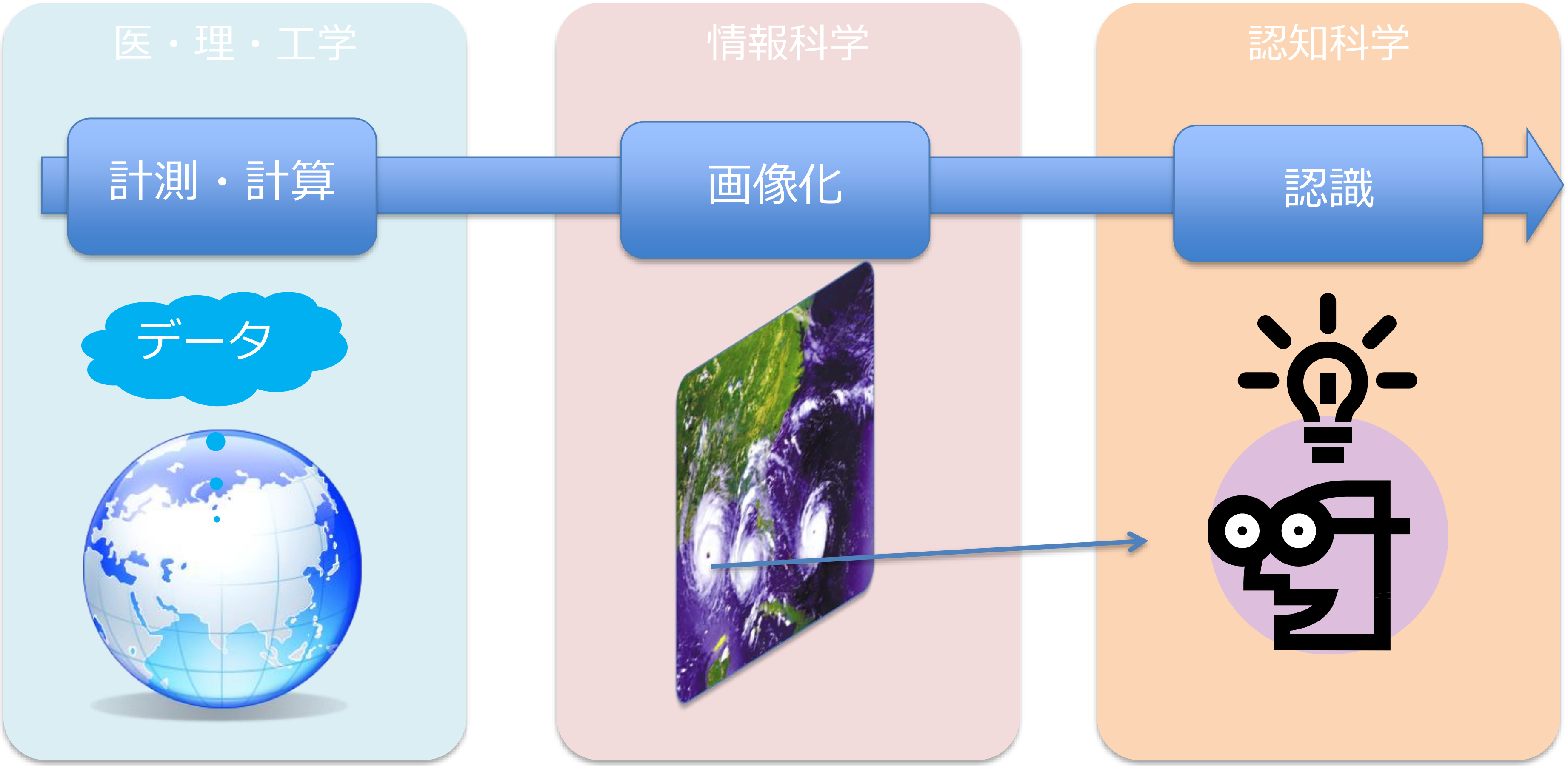
情報

人間が認識したデータ

可視化：データと脳をつなぐ



可視化研究と性能評価



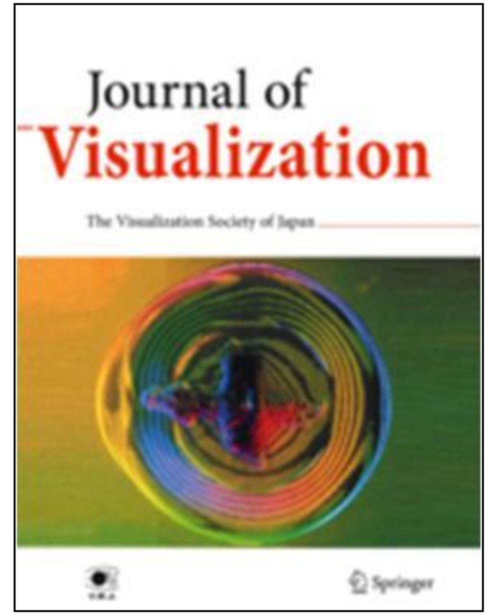
どんな現象をデータ化できたか？

どれほど効率よく画像化できたか？

どれほどの気付きを得たか、どんな行動変容に結び付いたか？

Journal of Visualization

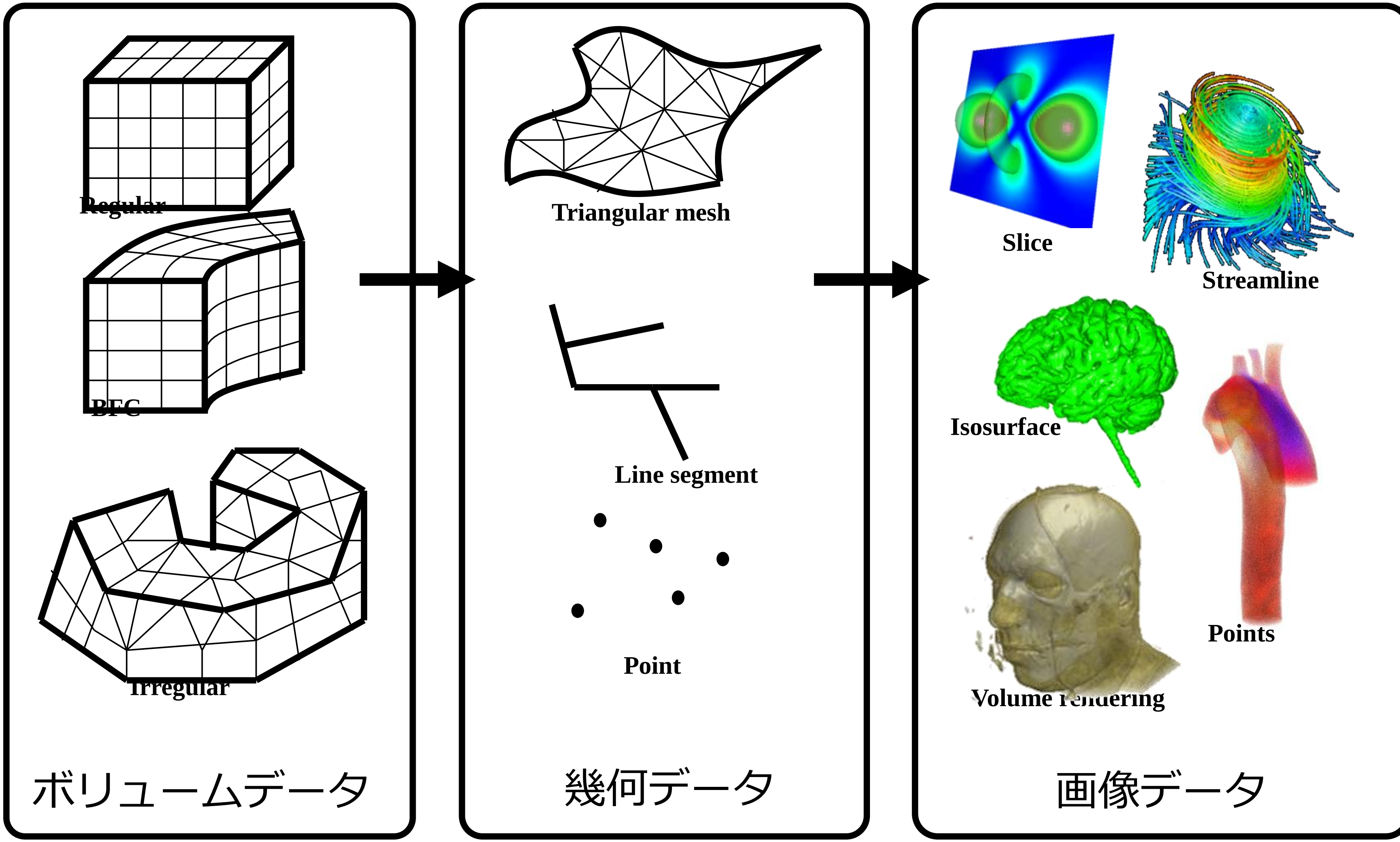
- History
  - Springer Vol.1(1998)-
- Editors-in-Chief:
  - K. Koyamada
  - K.C. Kim
- Scope
  - Visualization is an interdisciplinary imaging science devoted to making the invisible visible through the techniques of experimental visualization and computer-aided visualization.



データ可視化技術

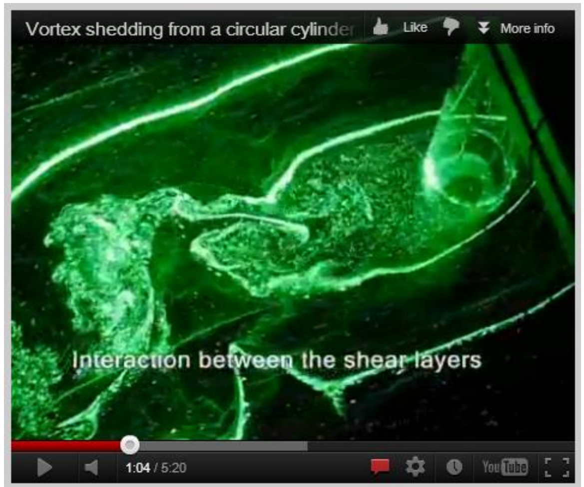
Visual analytics
ビジュアル分析

データ可視化パイプライン



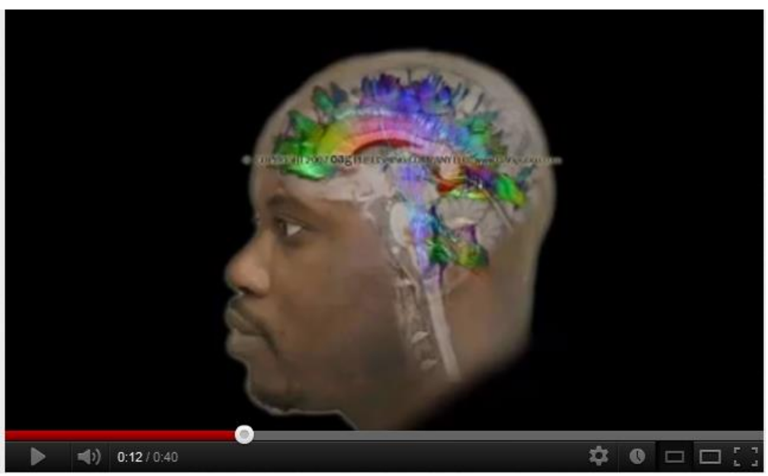
いろいろな可視化

計測データの可視化

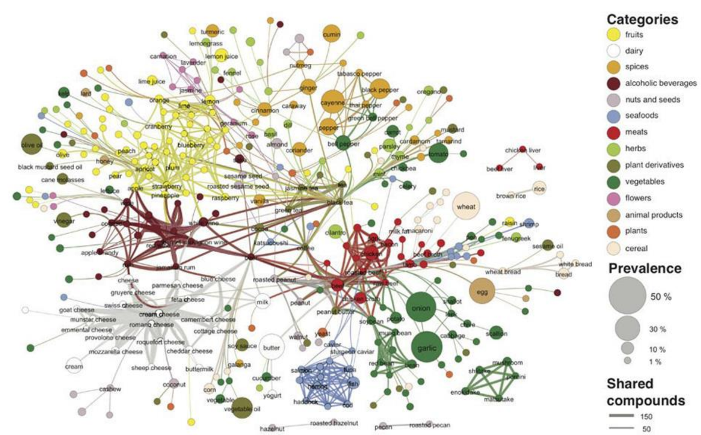


流れの可視化
（データ提供：Gustavo R.S. Assi）

船鉾の可視化
（データ提供：田中先生@立命館大学）

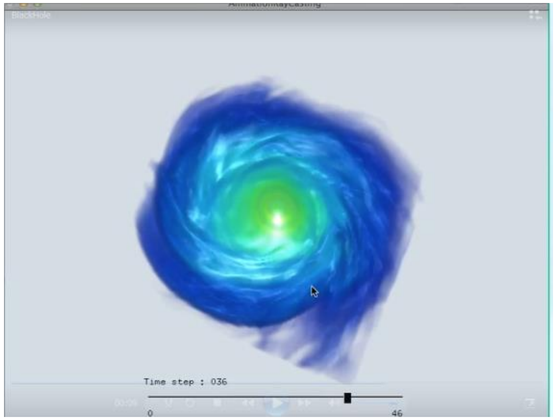


医療データの可視化
（データ提供：www.oagpubco.com）

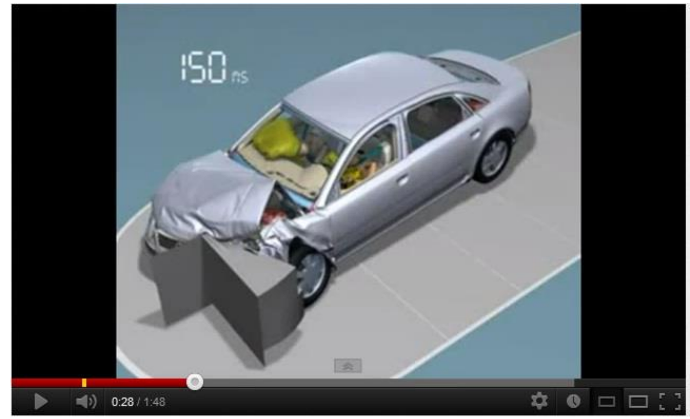


Yong-Yeol Ahn, et. al,
Flavor network and the principles of food pairing, Nature

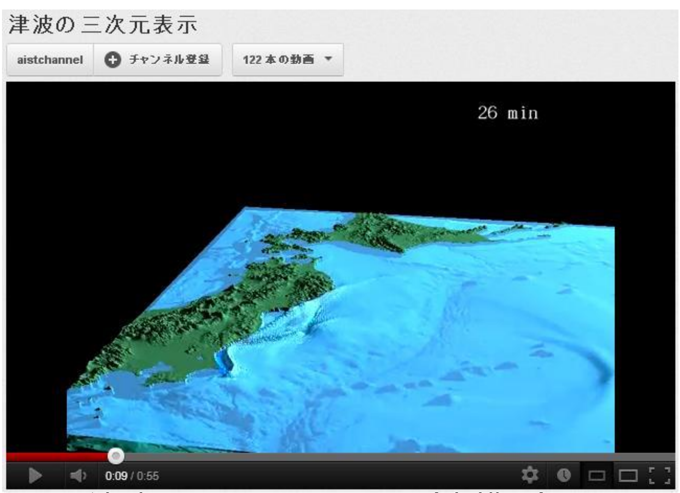
シミュレーション結果の可視化



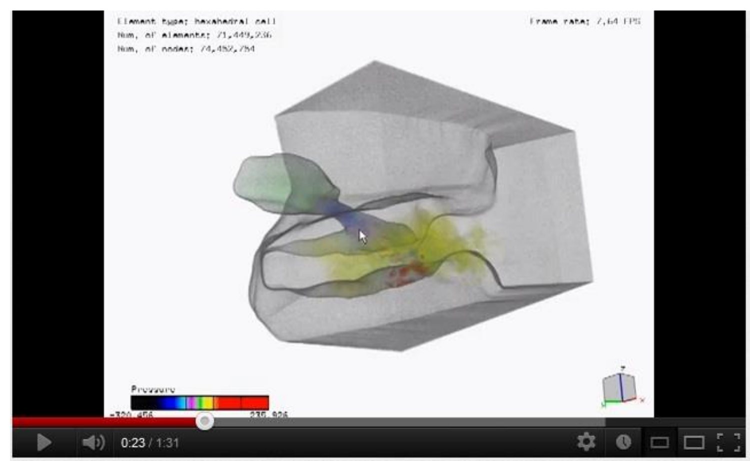
ブラックホールシミュレーション
（木内先生@京都大学）



自動車衝突解析(LOCTITE sitio Web www.loctite.com)

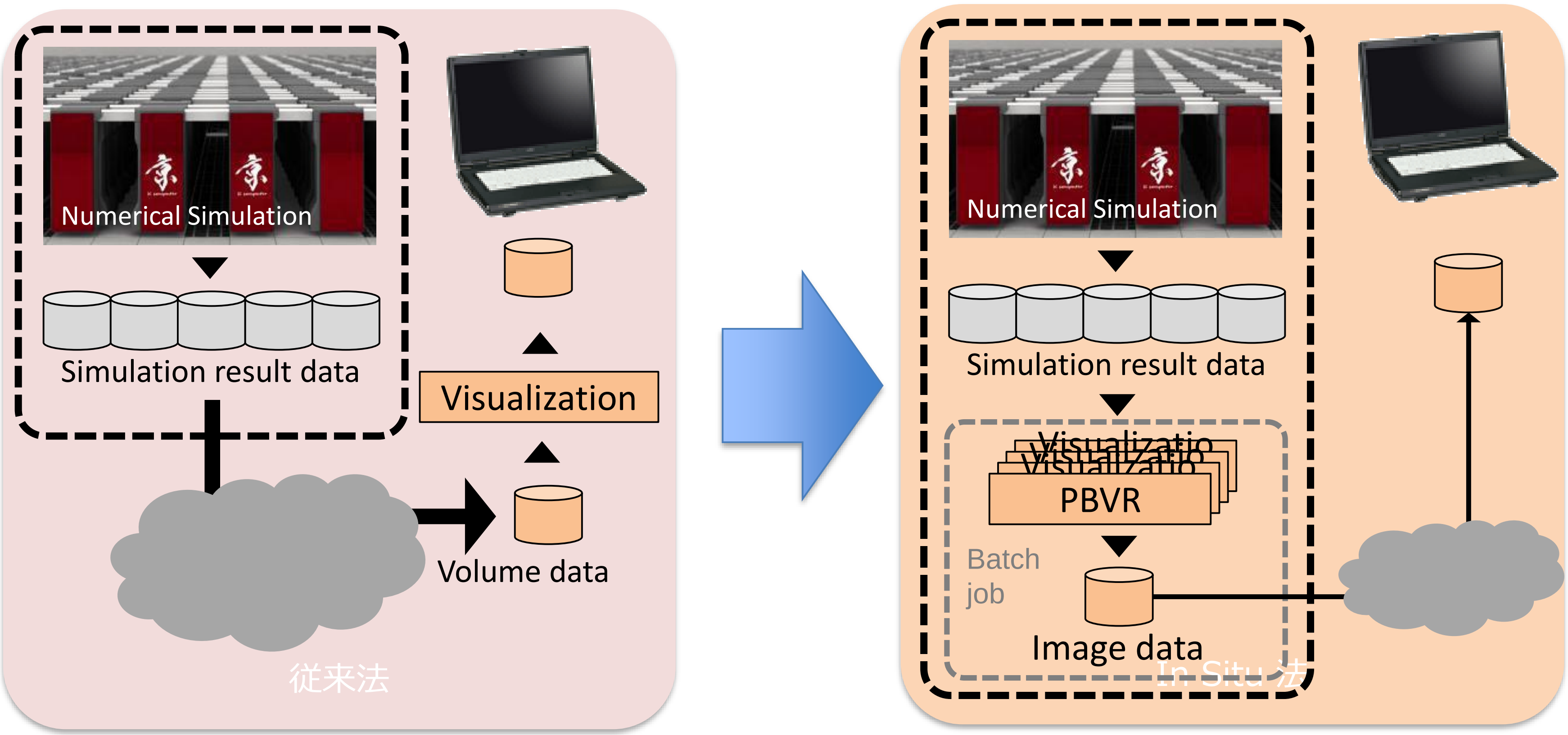


津波シミュレーション（産総研）



大規模口腔流体解析結果（野崎先生@大阪大学）

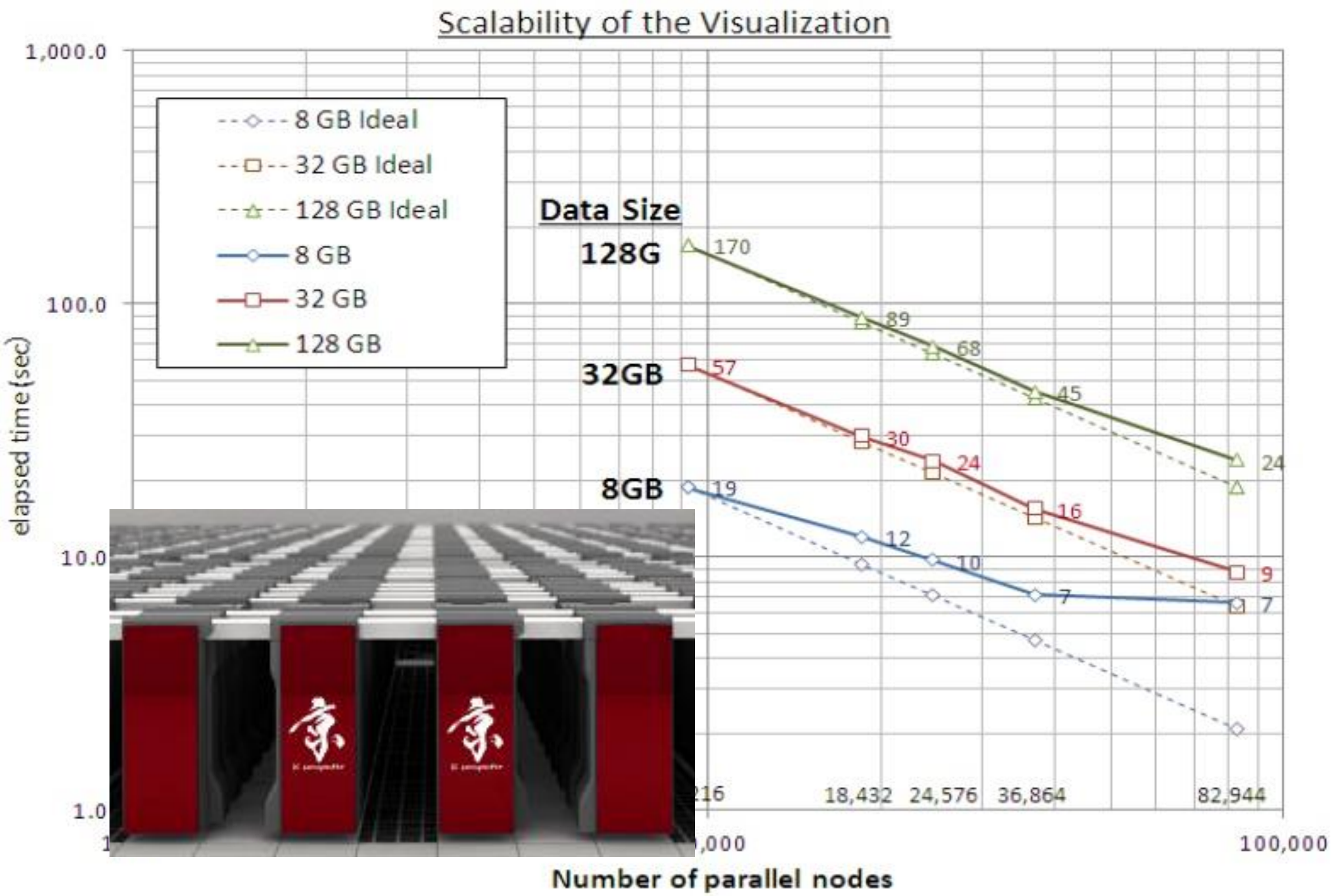
In-situ Visualization on K computer



経過時間：1時間以上　＞　経過時間：10秒程度

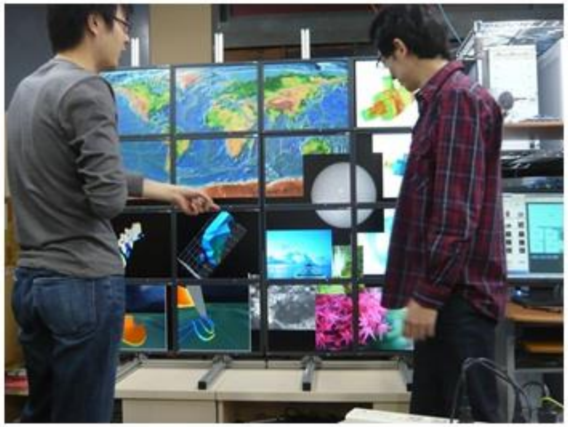
Ogasa et al., “Visualization technology for the K computer”, Fujitsu Scientific & Technical Journal, Vol.48, No.3, 2012.

Visualization in K-computer system



Ogasa, A; Maesaka, H; Sakamoto, K; Otagiri, S, “Visualization Technology for the K computer,” Fujitsu Sci. Tech. Journal, Vol.38 No.4, 2012

大型表示装置を使った可視化



複数アプリケーションの表示例

星間物質の乱流シミュレーション結果
（村主先生@白眉プロジェクト）

立体映像提示（アナグリフ方式）
（キッズサイエンススクール）

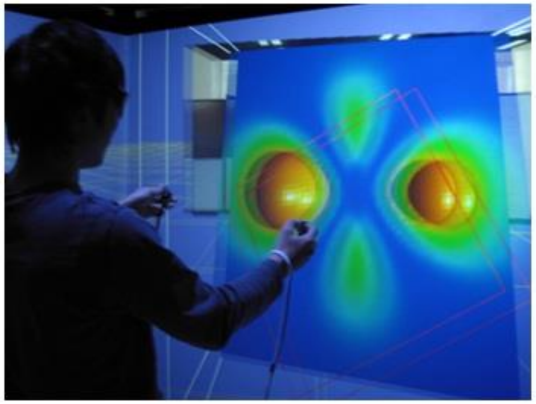
N体シミュレーション結果（暗黒物質）
（矢作先生@京大メディアセンター）

大規模ポンプの構造解析結果
（奥田先生@東京大学）

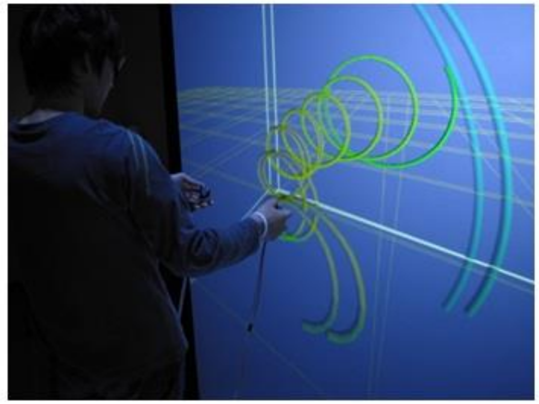
システム構成

PC x 11
NIC: 1Gbps
LCD: 40

没入表示装置を使った可視化

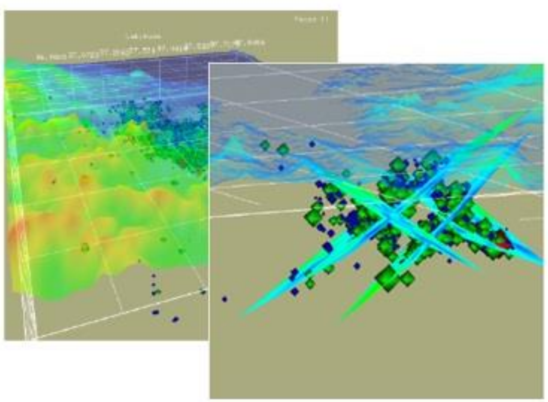


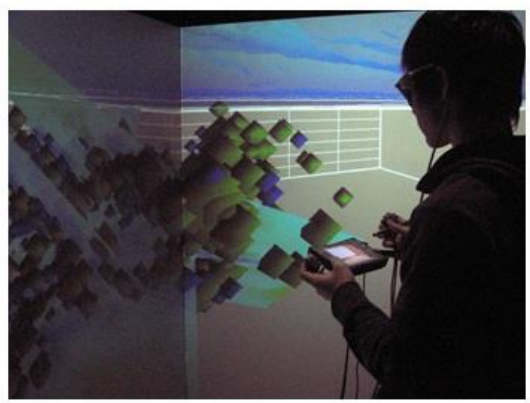
任意断面可視化



流線可視化

没入仮想空間提示





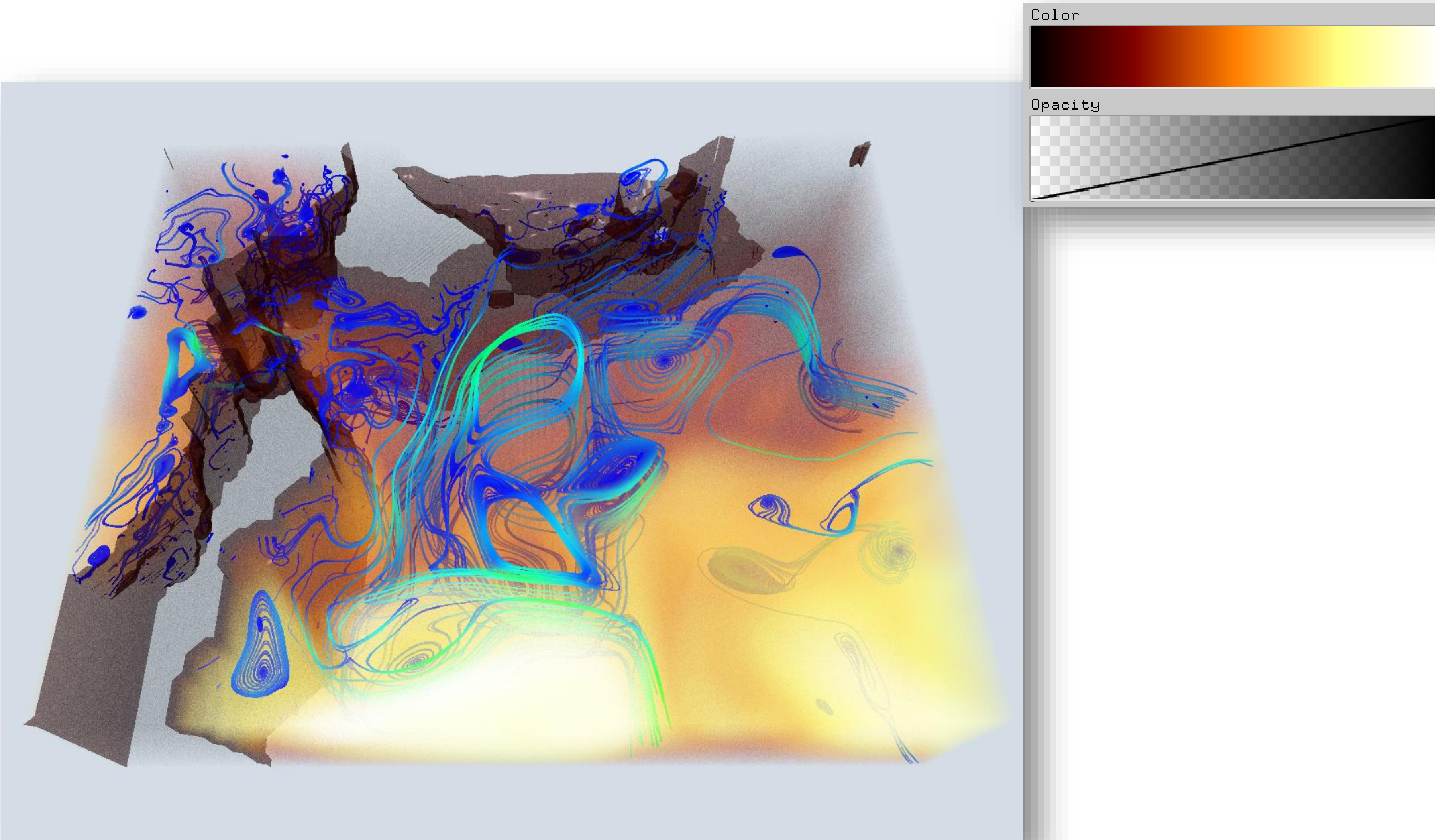
新潟中越地震の震源分布（断層面の推定）
（片尾先生@京大防災研）

- PC x 4
- 磁気装置
- 偏光メガネ
- 操作デバイス

システム構成

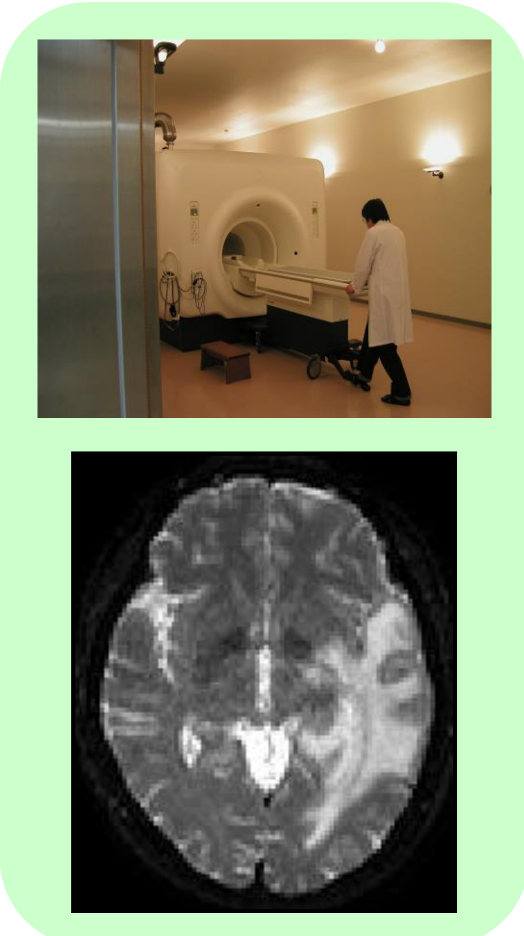
青森沖海流解析結果

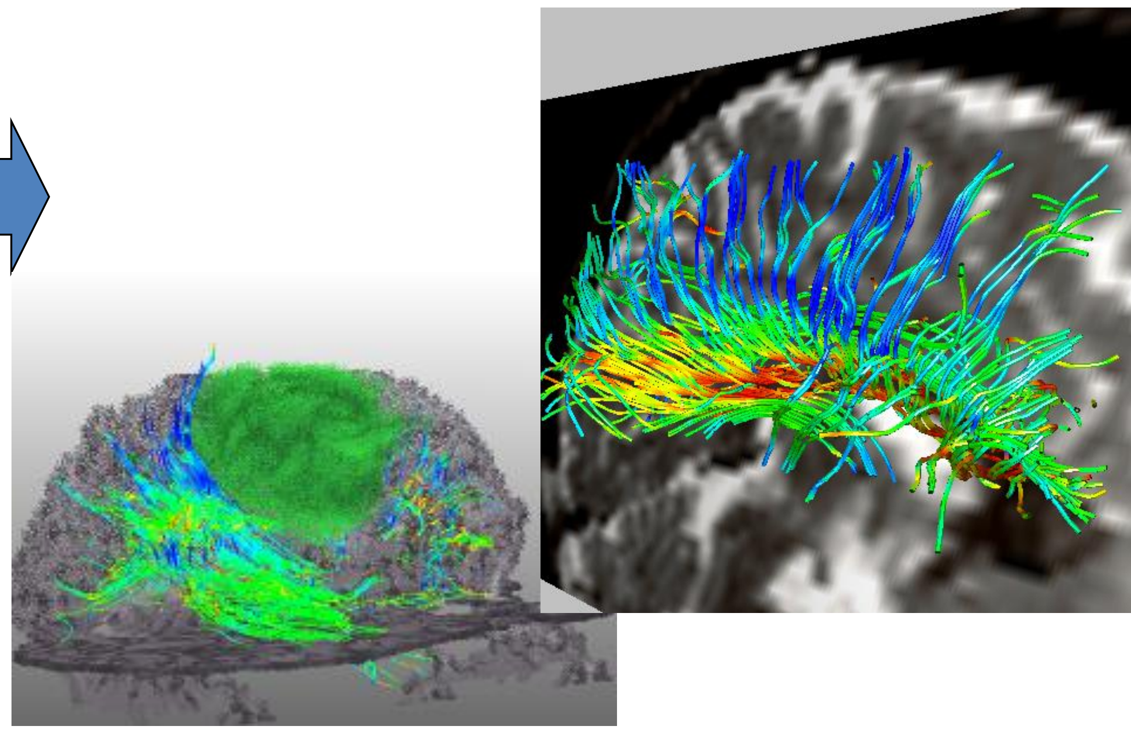
（石川先生@京大）



脳白質神経線維可視化技術

DT-MRIデータはテンソルデータの一種である．テンソルの最大固有値とその固有ベクトルから，神経線維の方向を計算し，流線可視化手法を適用することで，神経の流れを表現できる．



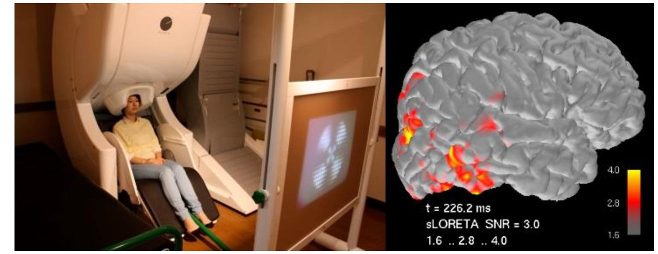


脳機能ネットワークの視覚的分析

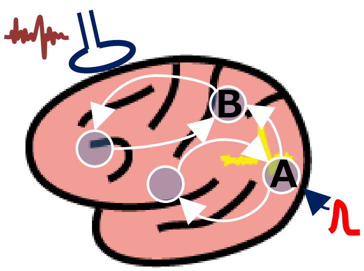
神経科学：「いつ」「どこで」脳が活動するか！

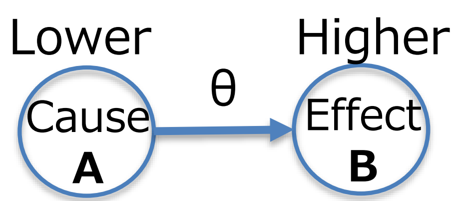
部位間の結合性、因果関係は未知な領域

脳はどのような、向きの有る結合（Network）を形成し、動的なシステムとして働くか？



MEGシステム　　脳機能イメージング

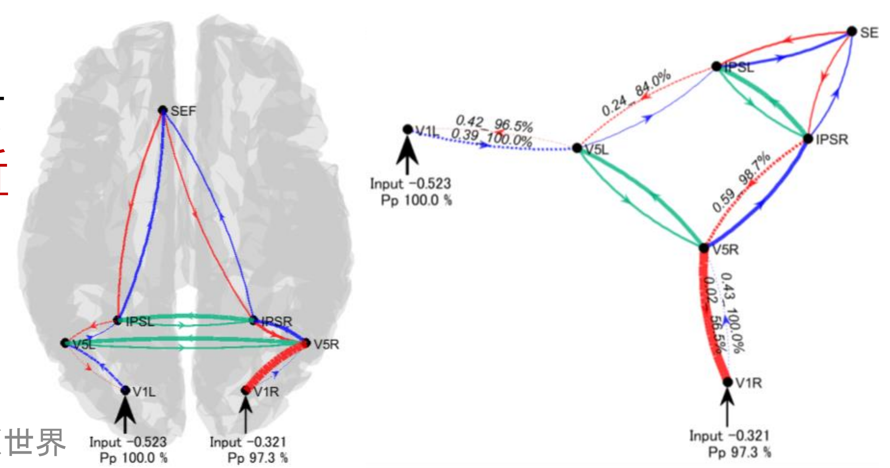




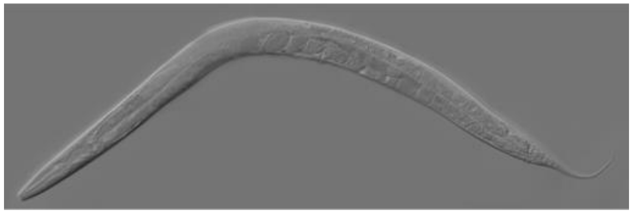
最新の因果推論技術と
- Granger Causality
- Convergent Cross Mapping
- Dynamical Causal Modeling

適切な可視化技術、複雑な解析手法に対して解析者への適切な推論を促すビジュアル分析を用いて、脳機能ネットワークを調べる

行動や認知に現れる疾患特異的な変化を脳神経ネットワークの変化として捉える

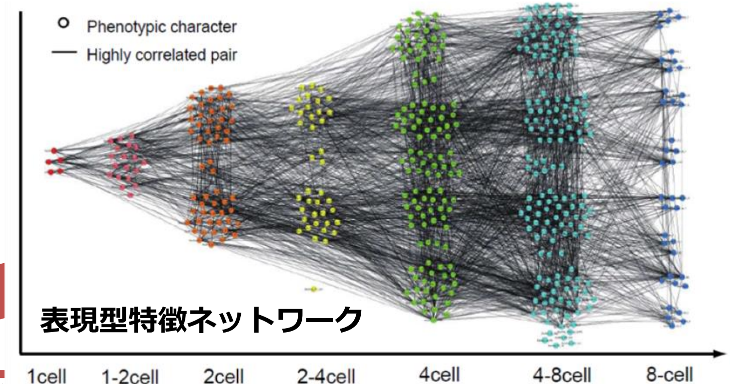


生命科学の視覚的分析



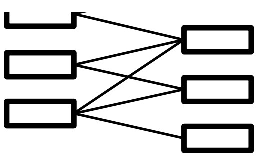
C. elegansの微分干渉顕微鏡像
(https://ja.wikipedia.org/wiki/カエノラブディティス・エレガンス)

バイオイメージインフォマティクス

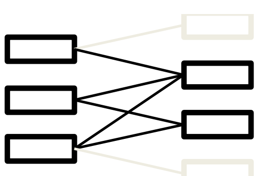


研究的問い：表現型特徴ネットワークと既知の生化学反応をどう結びつけるか？
仮説:ネットワークには陽的に表現されない潜在因子が存在する

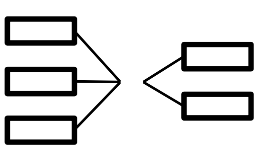
バイクラスタリングによるアプローチ



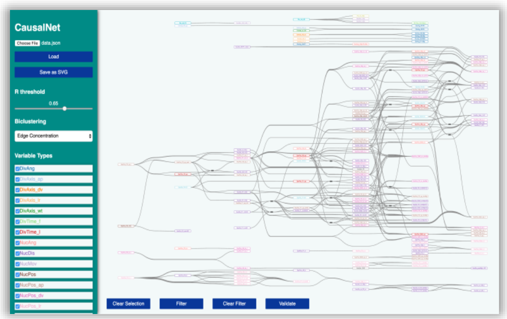
バイクラスタリング



潜在因子候補の抽出

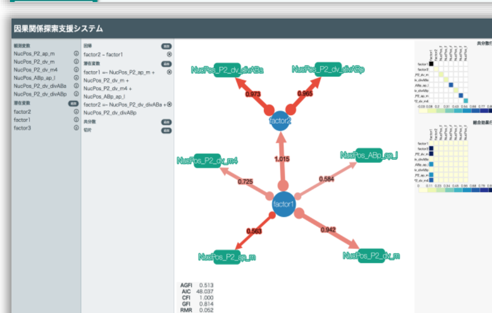


実装



仮説構築 - CausalNet

- 杉山フレームワーク、Edge Concentrationを用いた表現型特徴ネットワークの可視化
- バイクラスタリングによる潜在因子候補の抽出

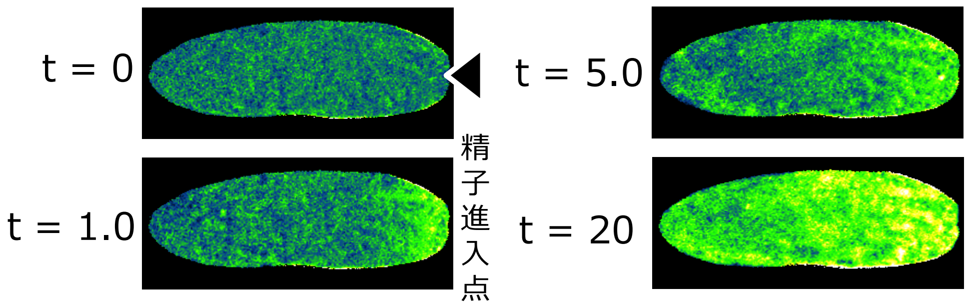


仮説検証 - iSEM

- 構造方程式モデリングを用いた因果関係、潜在変数の検証
- 対話的な統計モデル構築
- 統合的な統計モデル可視化

emo: https://youtu.be/Jltmd3zlQOE

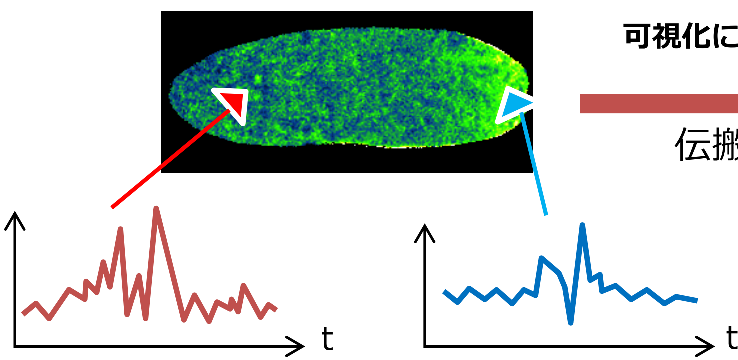
生命科学の視覚的分析



- 受精カルシウム波は
  生命共通の現象
- 線虫での解析手法の確立によって、ヒトの受精卵の理解に繋がる

研究的問い：遺伝子から一意に表現型が決まらないのはなぜ？

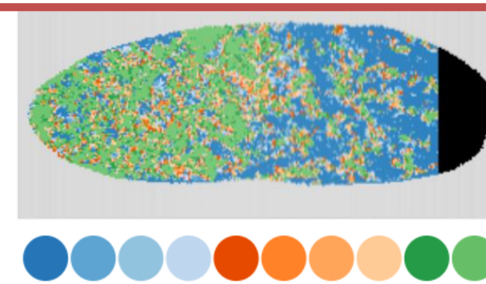
仮説：カルシウム波が表現型に影響を及ぼしている



従来の仮説：波は均一に一定速度で広がる

可視化による仮説検証

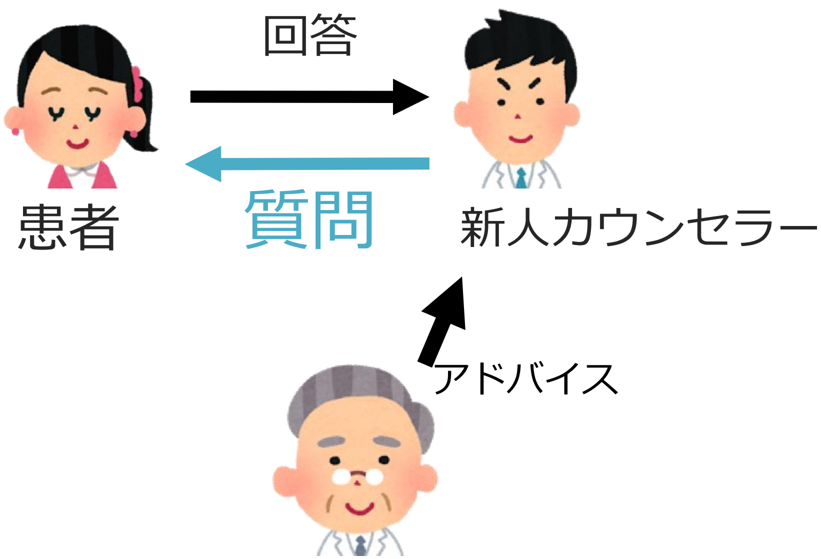
伝搬時間:0 1



可視化による発見：
波が均一に広がっていない可能性がある

会話の視覚的分析

背景：心療カウンセリング



熟練カウンセラーが後日、
質問内容に関してアドバイス。

現状では、熟練カウンセラーが
書き起こし文書を読んでから、
口頭でアドバイス。

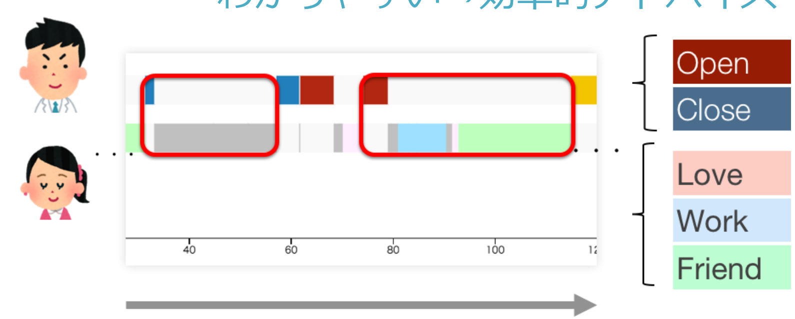
仮説：カウンセリングの会話の流れを可視化すれば、より効率的に
アドバイスできるのではないか

手法：

書き起こしテキストを入力 → 自然言語処理により発言を各カテゴリに分類 → 時系列に沿って可視化

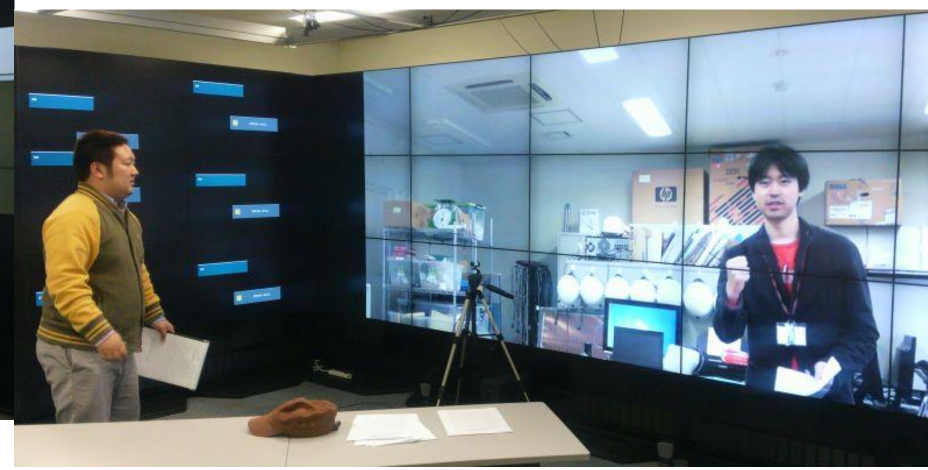
結果：

Ex:どのような質問からどのような回答が出たか
わかりやすい→効率的アドバイス



学習科学の視覚的分析

- ◆ Active Learning への効果的な支援が可能か？
  - ◆ 教材コンテンツ、発表資料提示
  - ◆ ビデオ会議システムと連携した遠隔授業への対応

◆現場の要求に対応した ユーザインタフェース技術が利用可能になれば、授業の円滑な進行を支援できる

◆ 授業での実践的活用を通した学習効果に関する評価・分析

可視化：ビッグデータ時代の科学を拓く

科学的可視化

子供の科学

原田三夫, “この雑誌の役目”, 1924年10月号

・・・しかしこの雑誌の一番大切な目的は、ほんとうの科学というものが、どういうものであるかを、皆さんに知っていただくことであります。ちかごろは「科学科学」とやかましくいいますが、ほんとうに科学というものを知っている人はたくさんないようです。人は生まれながら美しいものを好む心を持っていますが、それと同じように自然の物事について詳しく知り、深くきわめようとする欲があります。昔からその欲の強い人が調べた結果自然の物事の間には、たくさんの定まった規則のあることがわかりました。科学ということは、この規則を明らかにすることであります。多くの人が科学と言っているのは、たいていはその応用にすぎません。この規則を知ることによって人間は自然にしたがって無理のないように生き、楽しく暮らすことができ、これを応用して世が文明におもむくのです。

科学とは、物事の間の因果関係を明らかにすること

科学する心

小林秀雄, “科学する心”, 岩波書店講演集CD, 1970年

・・・それは、まぁこの環境を知る一種の科学論だなぁ。科学論てものはねぇ、とてももう面倒なものなんですよ。で、ただ、科学ってものはだなぁ、ものがその、ものをこの本当にものをしるのが科学ではない。あれはものの法則を知るんです。いいですか、そこはねぇ、あのよおく考え貰わないといけないんだよ。つまりだなぁ、えー、科学ってものはいつでも法則をめがけているんです。ほんとに僕らの経験ってものを、ほんとの生きる経験ってものは科学はいらないんです。生きていなくてもいいんです。生きている人間ってものは、科学は、そんなものは認めないんです。いつでも、科学はねぇ、その、規則をめがけるんです。ものとものはどういう関係にあるかってことを、因果関係だね、一口に言えば因果関係ってものは、自然はどういうふうに動いているかっと、いうその因果関係っていうものを目指しているんです・・・

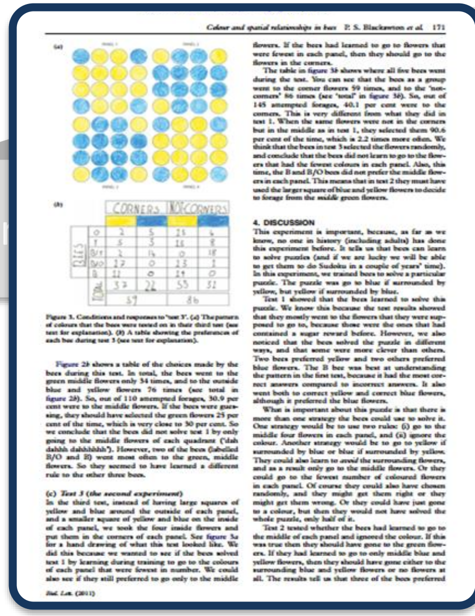
科学は因果関係を目指す

科学教育

biology letters

Animal behaviour

Blackawton bees



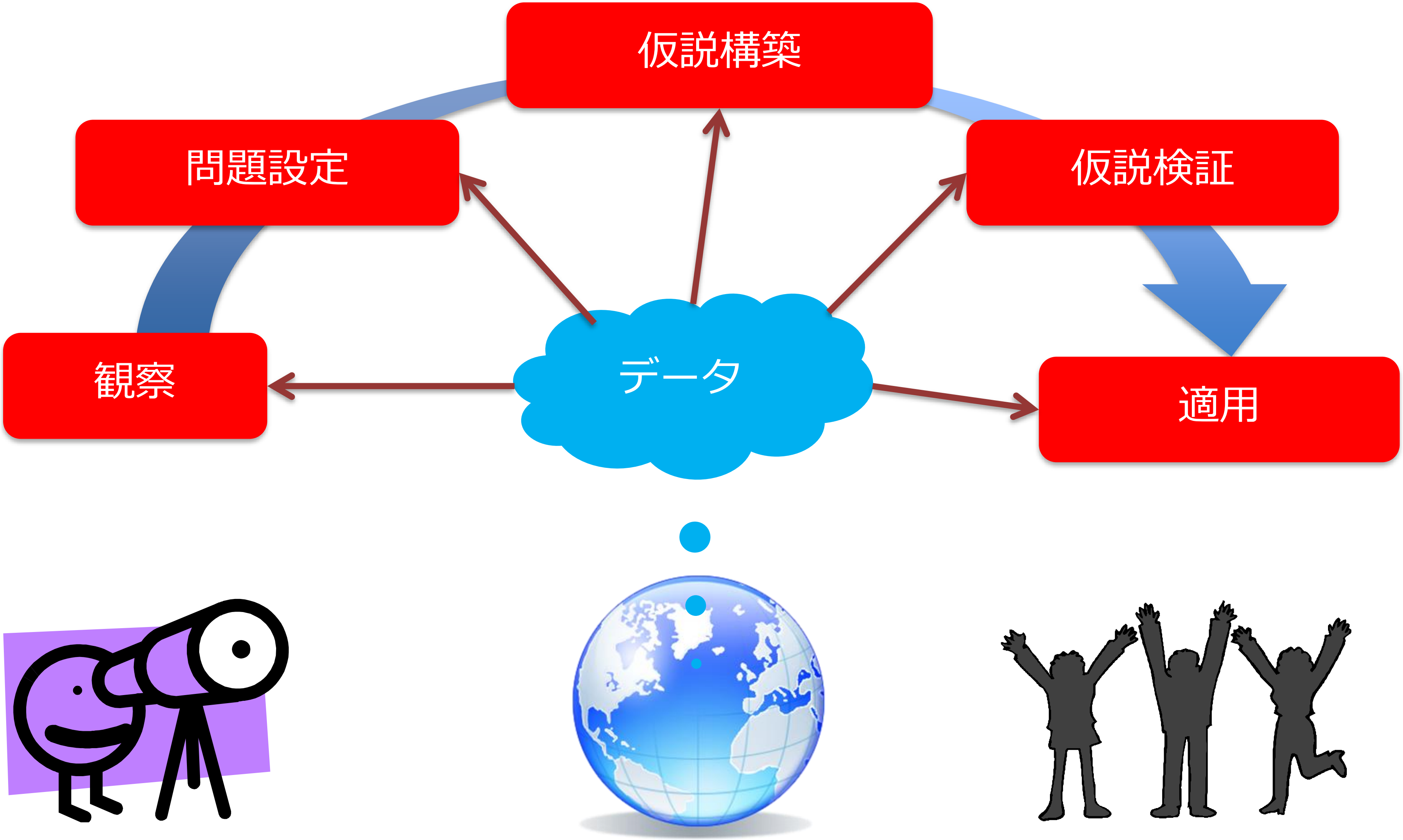
Colour and spatial relationships in bees

4. DISCUSSION

Elementary school students,
"Blackawton bees,"
Biology Letters
(2011) 7, 168–172

- USやEUでは, 小学校で科学的方法が教育されている
- 日本では, 大学や大学院で卒論や修論を通じて暗黙的に伝えられる
  - 日本では科目試験（正解がひとつ）に極端にこだわる
  - ほとんどの日本人学生は科学的方法をキチンと学ばない

科学的方法の流れ



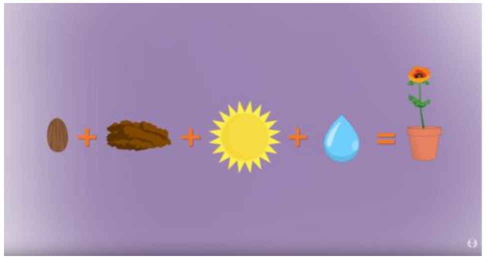
科学的方法

https://www.youtube.com/watch?v=KIFz_-KzURY

Scientific Method Song Video

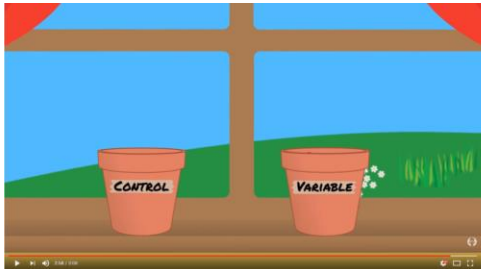
Make an observation

Ask a question



Form a hypothesis

Make a prediction



Do a test or experimentation

Analyze data and draw a conclusion

「科学的方法」を一緒に体験してみよう！

- 前提
  - 花の成長に興味がある
  - 花の成長の指標として、花の重さを考える
  - 花の成長を説明する変数として、養分量、日照時間、水の量を考える
- 質問
  - 花の成長を説明するためにどのような実験を行うのがよいかな？

実験を計画しよう

- 測ろう(被説明変数):花の重さ
- 変化させよう(説明変数):与えた水の量
- 同じようにしておこう(制御変数):日照時間、養分量

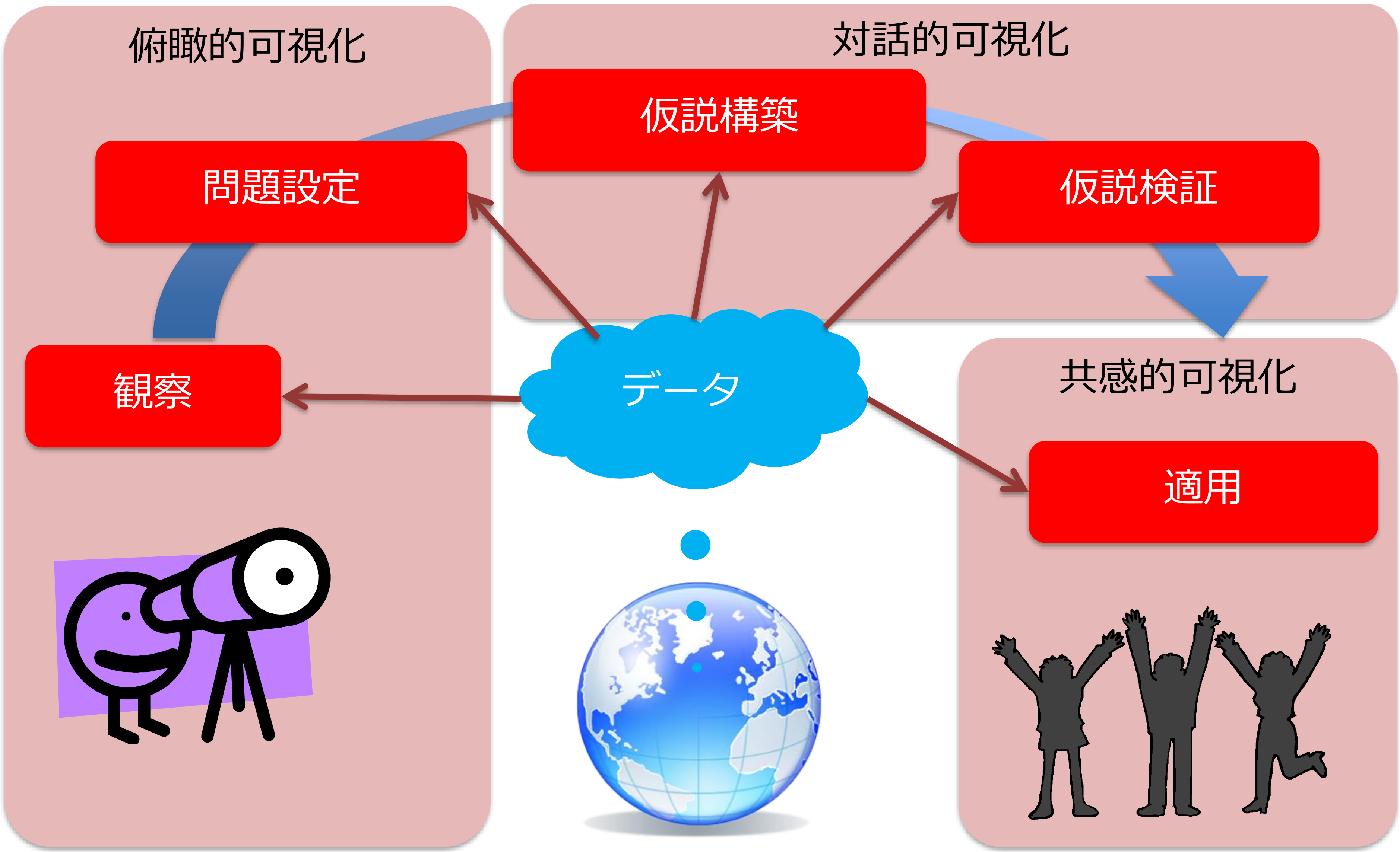
鉢番号	水の量	花の重さ
1	46	16
2	69	24
3	90	28
4	13	8
5	54	19
6	46	15
7	53	19
8	35	12

実験を計画しよう

- 被説明変数（従属変数とも）：花の重さ
- 変化させよう（独立変数とも）：与えた水の量
- 同じようにしておこう（制御変数）：日照時間、養分量

鉢番号	養分量	水の量	花の重さ
1	2	3	19
2	4	1	40
3	1	2	10
4	2	1	19
5	4	3	32
6	4	2	40
7	2	4	13
8	4	2	34

科学的方法と可視化



可視化：ビッグデータ時代の科学を拓く

科学と因果発見技術

仮説

- ふたつの変数間のある方向をもった仮の関係
- その関係は、観測データによって検証される
- 「もしXならばY」のような因果関係によって記述される

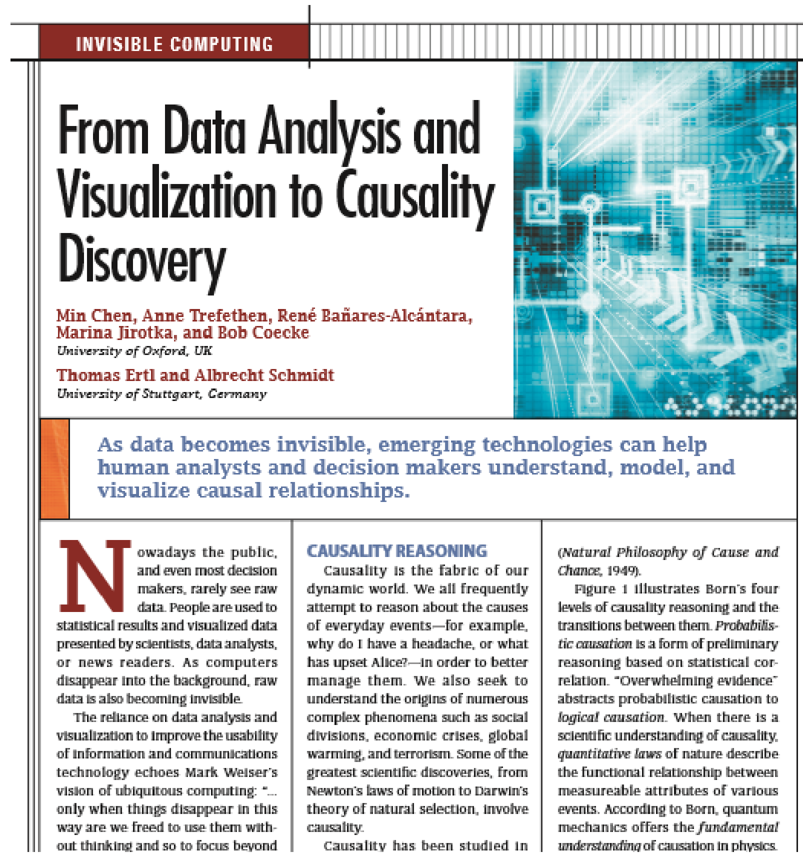
因果関係発見は、科学における本質である

因果発見技術

M. Chen, et al., “From data analysis and visualization to causality discovery,” IEEE Computer , 44(10):84-87,

四つの挑戦

- 因果推論を促進するためのインフラ整備
- 可視化活用の能力の飛躍的向上
- 継続的なデータ洪水の対話的処理における負荷の軽減
- 推論過程における不確からしさの評価・可視化の対する要求への対応


INVISIBLE COMPUTING

From Data Analysis and Visualization to Causality Discovery

Min Chen, Anne Trefethen, René Bañares-Alcántara, Marina Jirotka, and Bob Coecke
University of Oxford, UK
Thomas Ertl and Albrecht Schmidt
University of Stuttgart, Germany

As data becomes invisible, emerging technologies can help human analysts and decision makers understand, model, and visualize causal relationships.

Nowadays the public, and even most decision makers, rarely see raw data. People are used to statistical results and visualized data presented by scientists, data analysts, or news readers. As computers disappear into the background, raw data is also becoming invisible.

The reliance on data analysis and visualization to improve the usability of information and communications technology echoes Mark Weiser's vision of ubiquitous computing: "... only when things disappear in this way are we freed to use them without thinking and so to focus beyond

CAUSALITY REASONING

Causality is the fabric of our dynamic world. We all frequently attempt to reason about the causes of everyday events—for example, why do I have a headache, or what has upset Alice?—in order to better manage them. We also seek to understand the origins of numerous complex phenomena such as social divisions, economic crises, global warming, and terrorism. Some of the greatest scientific discoveries, from Newton's laws of motion to Darwin's theory of natural selection, involve causality.

Causality has been studied in

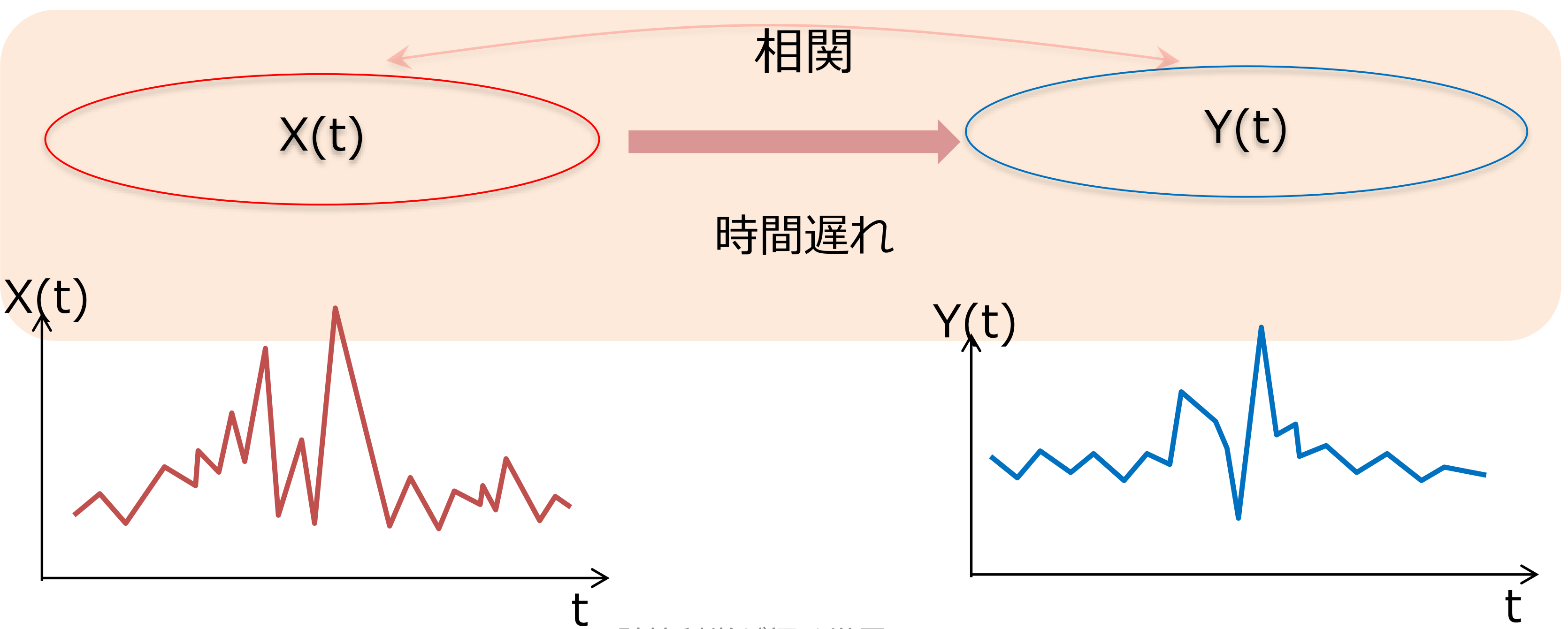
(*Natural Philosophy of Cause and Chance*, 1949).

Figure 1 illustrates Born's four levels of causality reasoning and the transitions between them. *Probabilistic causation* is a form of preliminary reasoning based on statistical correlation. "Overwhelming evidence" abstracts probabilistic causation to *logical causation*. When there is a scientific understanding of causality, *quantitative laws* of nature describe the functional relationship between measureable attributes of various events. According to Born, quantum mechanics offers the *fundamental understanding* of causation in physics.

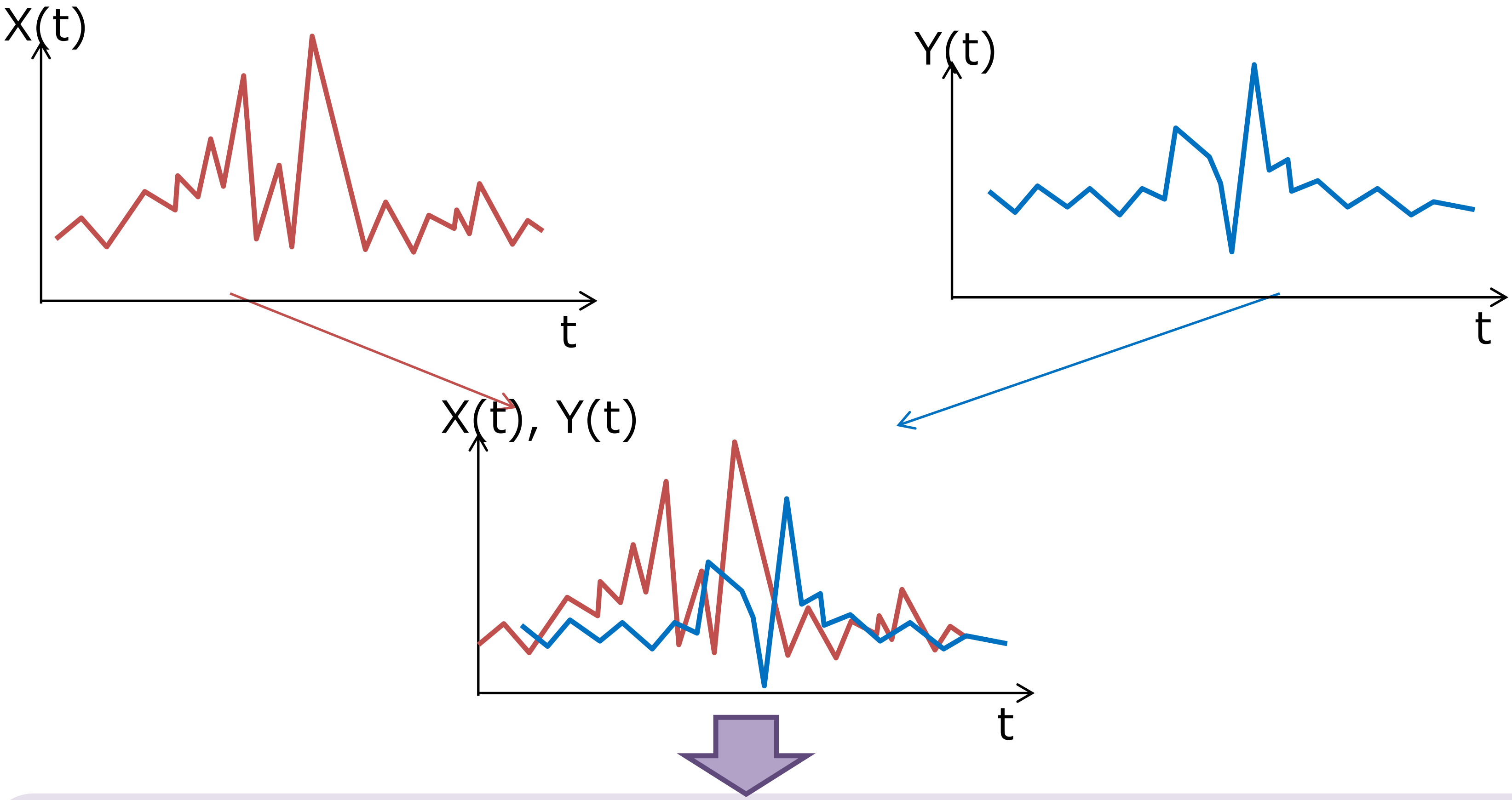
よりよい理解に向けて

時系列データ向け因果発見

- 因果関係は、あるプロセスともう一つのプロセスとの間の関係を表す
- 最初のプロセスは、二番目のプロセスに対して、部分的に関与する。
- 統一された定義はないものの、さまざまな因果発見技術が報告され、実際に活用されている。



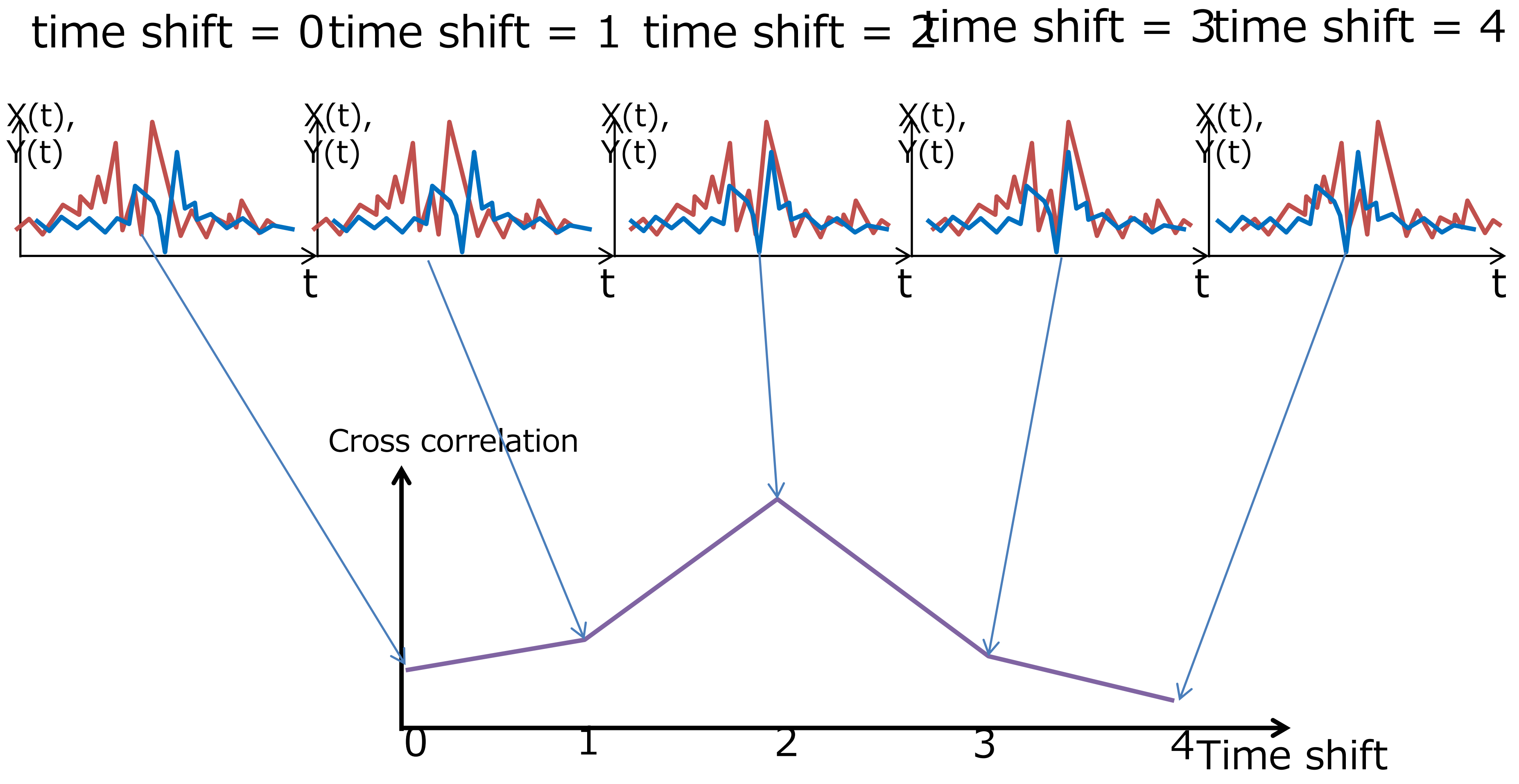
因果発見のための統計的検証法



- 相互相関
- グレンジャー因果
- 収束的交差写像

因果確率

相互相関



time shift = 0
time shift = 1
time shift = 2
time shift = 3
time shift = 4
X(t),
Y(t)
t
Cross correlation
0
1
2
3
4
Time shift

グレンジャー因果(GC： Granger Causality)

Granger, C. W. J. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods". Econometrica 37 (3): 424–438 (1969)

It is a statistical test for a cause-and-effect relationship between two time series variables

We perform two vector auto-regressions as follows:

$$Y(t) = \sum_{l=1}^{L} a_l Y(t-l) + \epsilon_1 \quad (1)$$

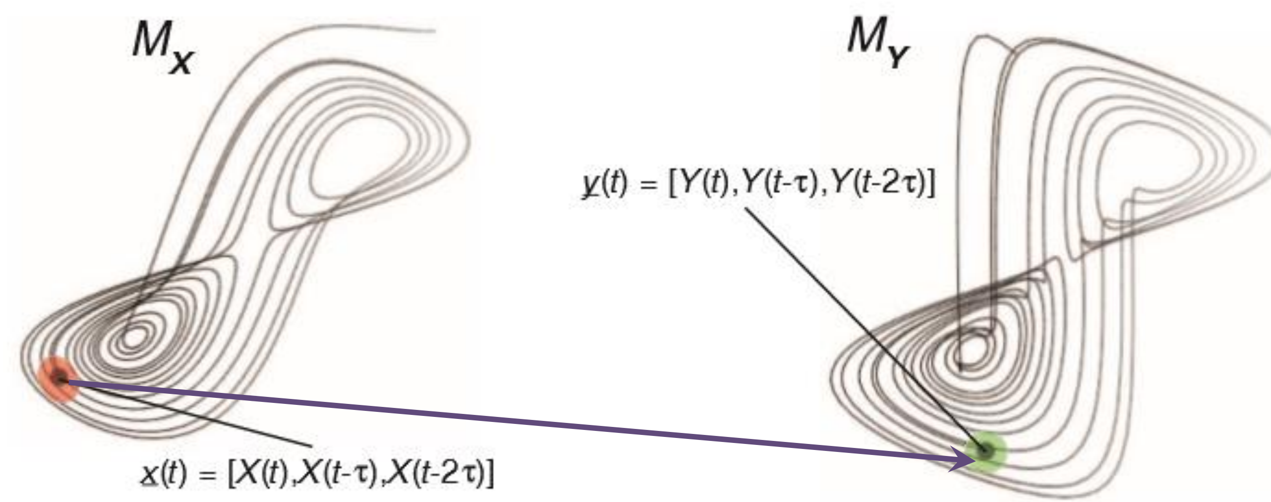
$$Y(t) = \sum_{l=1}^{L} a_l' Y(t-l) + \sum_{l=1}^{L} b_l' X(t-l) + \epsilon_2, \quad (2)$$

where L is the maximal time lag. We say X causes Y if eq (2) is statistically significantly better than eq (1).

収束的交差写像(CCM: Convergent Cross Mapping)

Sugihara, George; et al., "Detecting Causality in Complex Ecosystems," Science 338 (6106) pp. 496–500 (2012)

- It is a statistical test for a cause-and-effect relationship between two time series variables
- It constructs two shadow manifolds, M_X and M_Y using lagged-coordinate embedding X and Y, respectively
- If X forces Y uni-directionally, variable Y will contain information about X, but not vice versa.

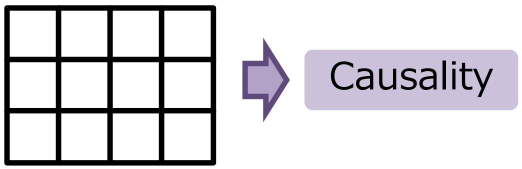


GC と CCM 使い分け

- どちらも相関は必ずしも因果を意味しないという問題を解決しようとする
- GCは、二つの変数の影響がお互いに独立で分離可能であるような確率的システムに適する
- CCMは、動的システム理論に基づき、二つの変数が相乗効果をもつようなシステムに適する

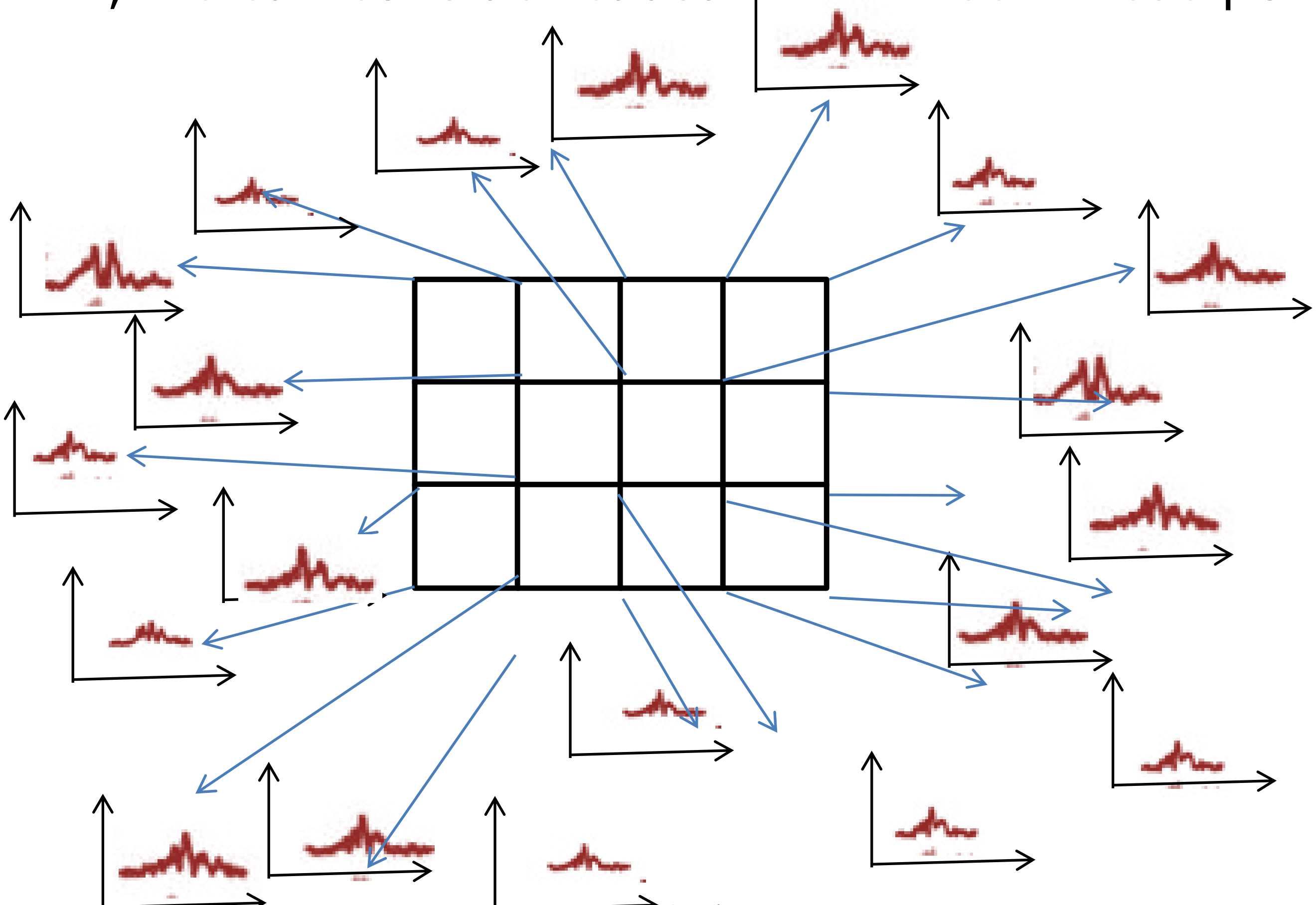
可視化：ビッグデータ時代の科学を拓く

時系列ボリュームデータにおける因果関係

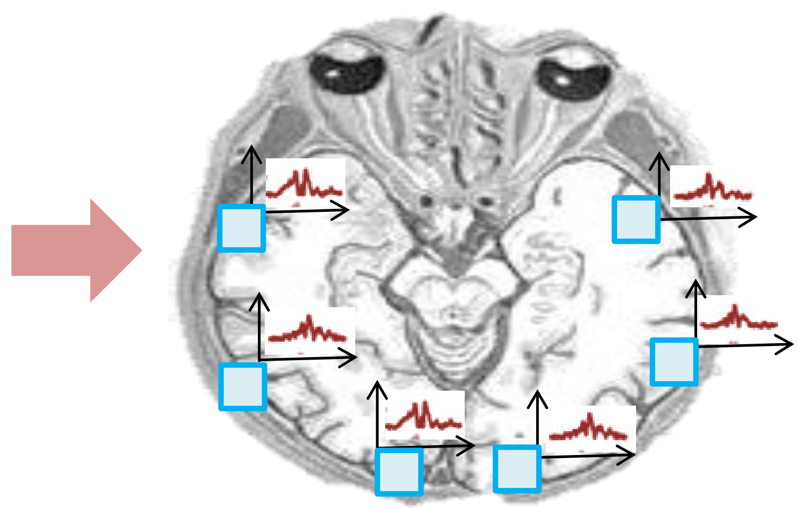


単一変数の時系列ボリュームデータ

- It defines univariate time series data defined at each grid point
- With it, we can consider causality between these points



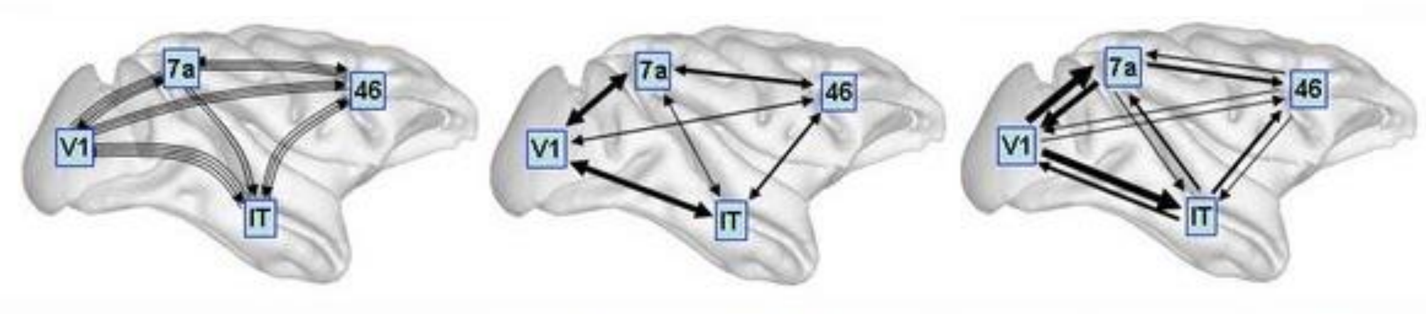
脳磁図(MEG: Magnetoencephalography)



- It　records magnetic fields produced by electrical currents occurring naturally in the brain, using very sensitive magnetometers.
- Its applications include basic research into perceptual and cognitive brain processes, localizing regions affected by pathology before surgical removal, determining the function of various parts of the brain, and neurofeedback.

結合解析

structural connectivity　functional connectivity　effective connectivity

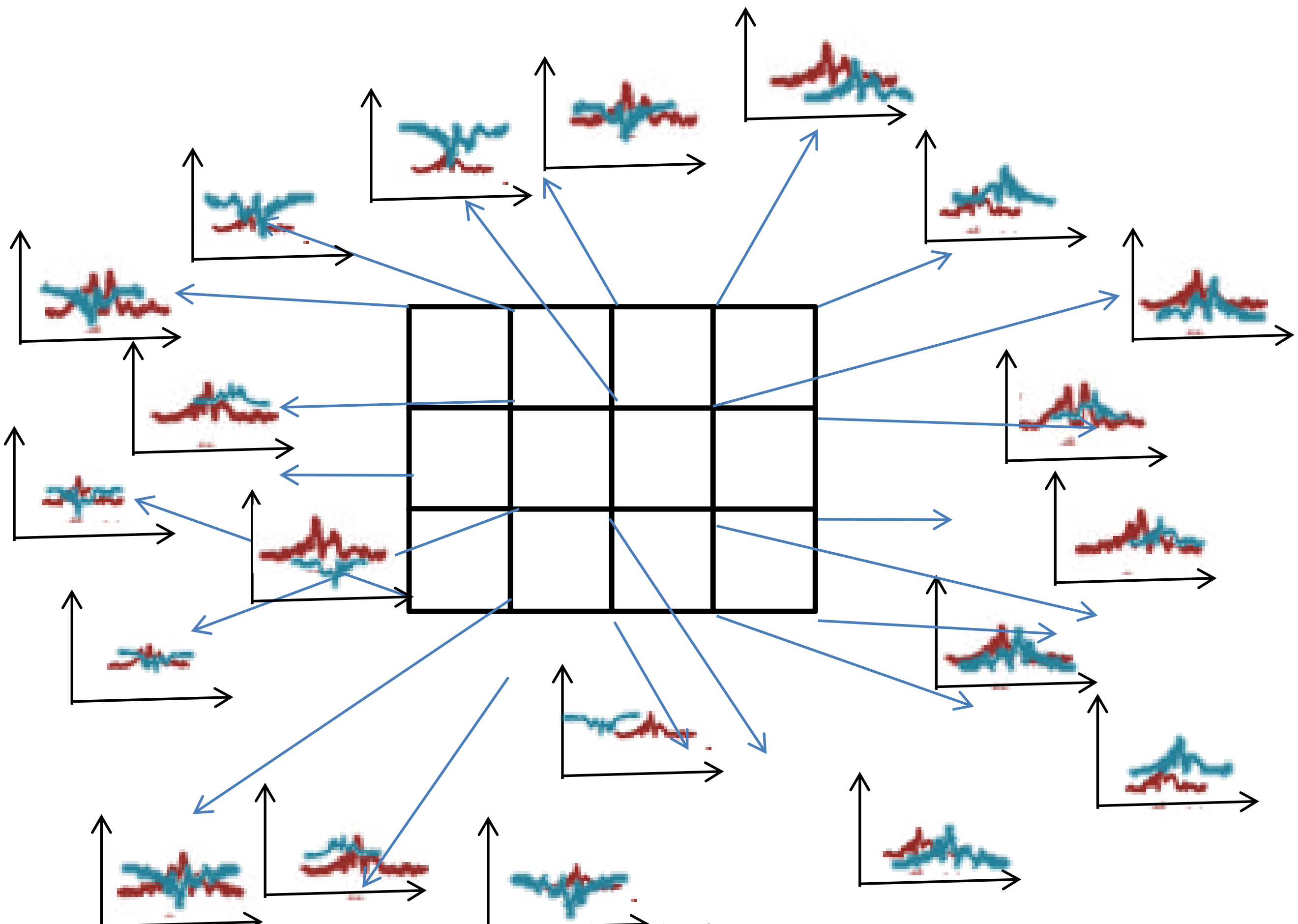


O. Sporns 2007, *Scholarpedia*

- Evaluation of structural connectivity (anatomical) using MR-DTI and Fiber-Tracking
- Confirmation of functional connectivity between activated areas using correlation
- Analysis of effective connectivity in neuronal groups using causation

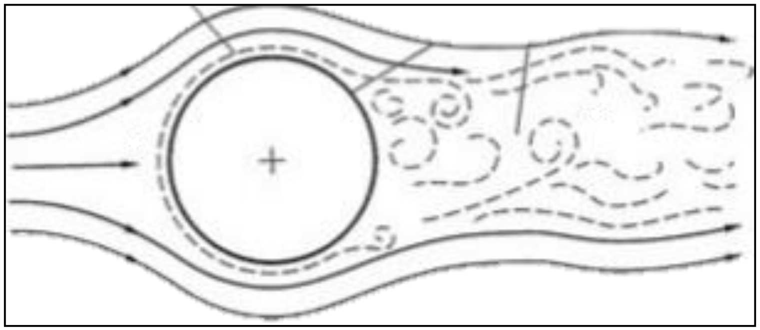
二変数の時系列ボリュームデータ

- It defines bivariate time series data defined at each grid point
- With it, we can consider causality at the point



洗濯効果

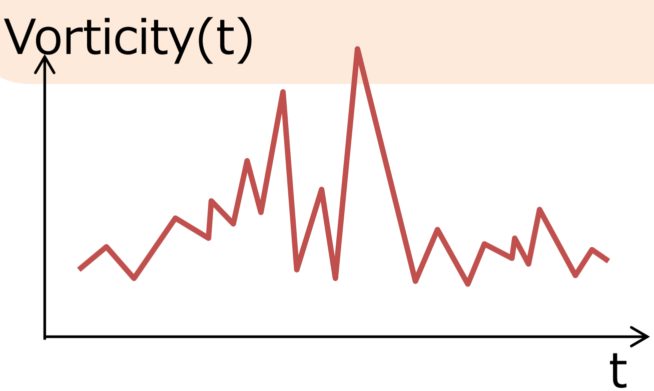
K. Suzuki, “UNSTEADY HEAT TRANSFER IN A CHANNEL OBSTRUCTED BY AN IMMERSED BODY”, 1994

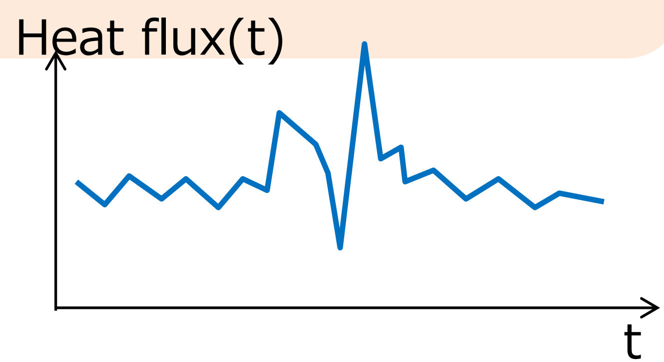


The vortex flow contributes to the enhancement of the heat transfer

Causality ?

Vorticity → Heat flux

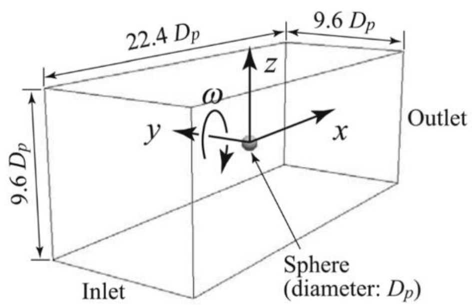


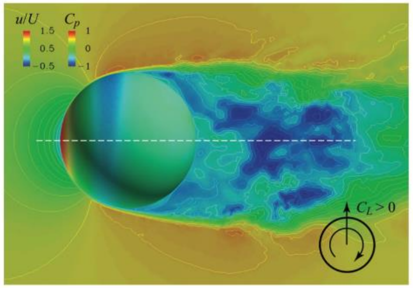


熱球周りの流体解析

Confirmation of the washing effect in the separated flow in the wake behind a sphere

- Model description
  - Number of nodes： 15,321,546
  - Number of cells： 18,917,887 (prism)
    27,545,304 (tet)
  - Number of time steps：198
  - File size：3.5GB per step

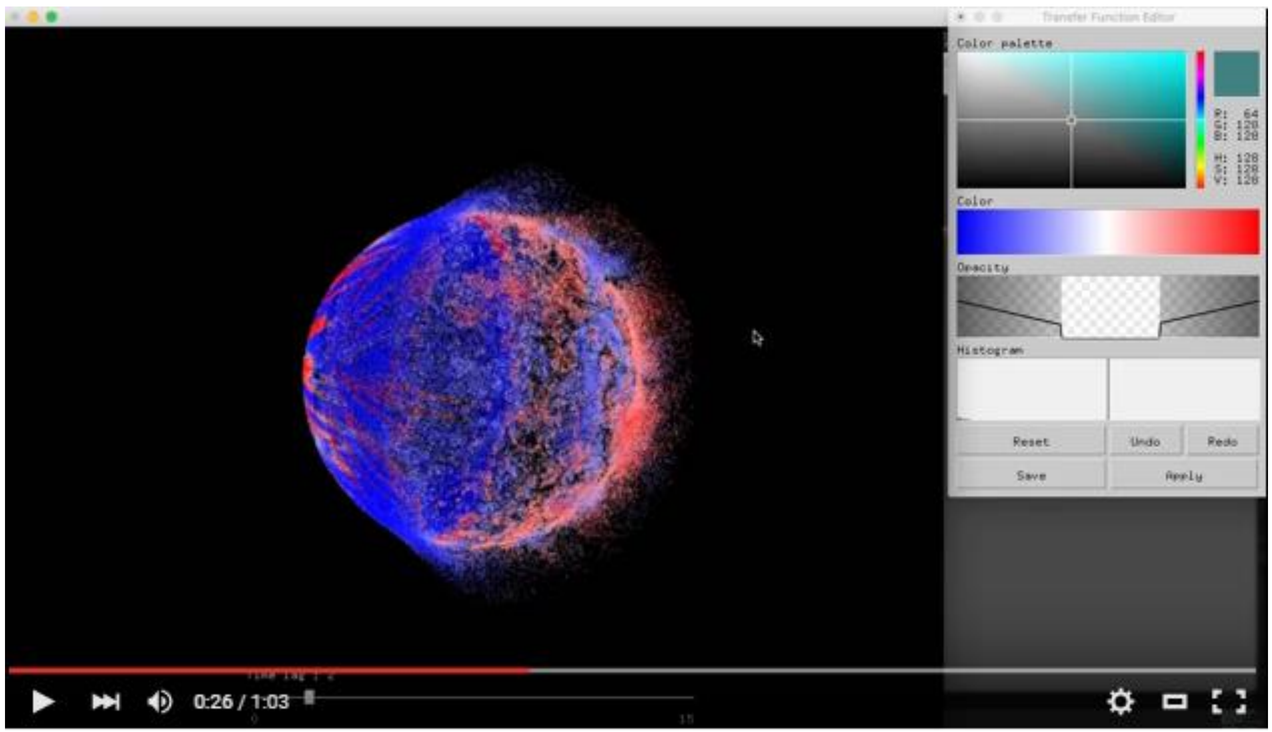




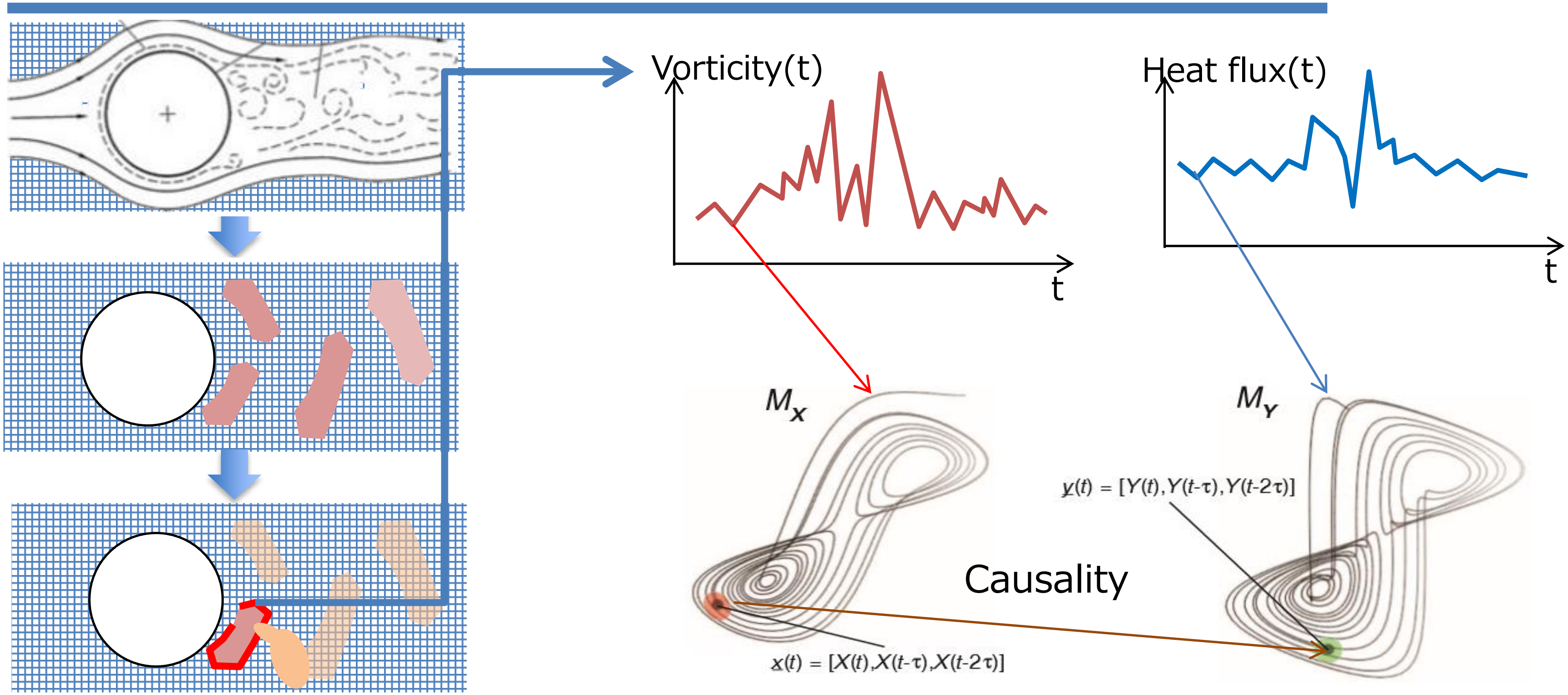
Negative Magnus lift on a rotating sphere at around the critical Reynolds number
Muto, Tsubokura, Oshima, Physics of Fluids, 24, 014102（2012）

因果関係ボリュームデータ

- Calculation of cross correlation at each grid point by changing the shift time
- Creation of a time-varying volume with 16 shift time steps



因果関係発見に向けて

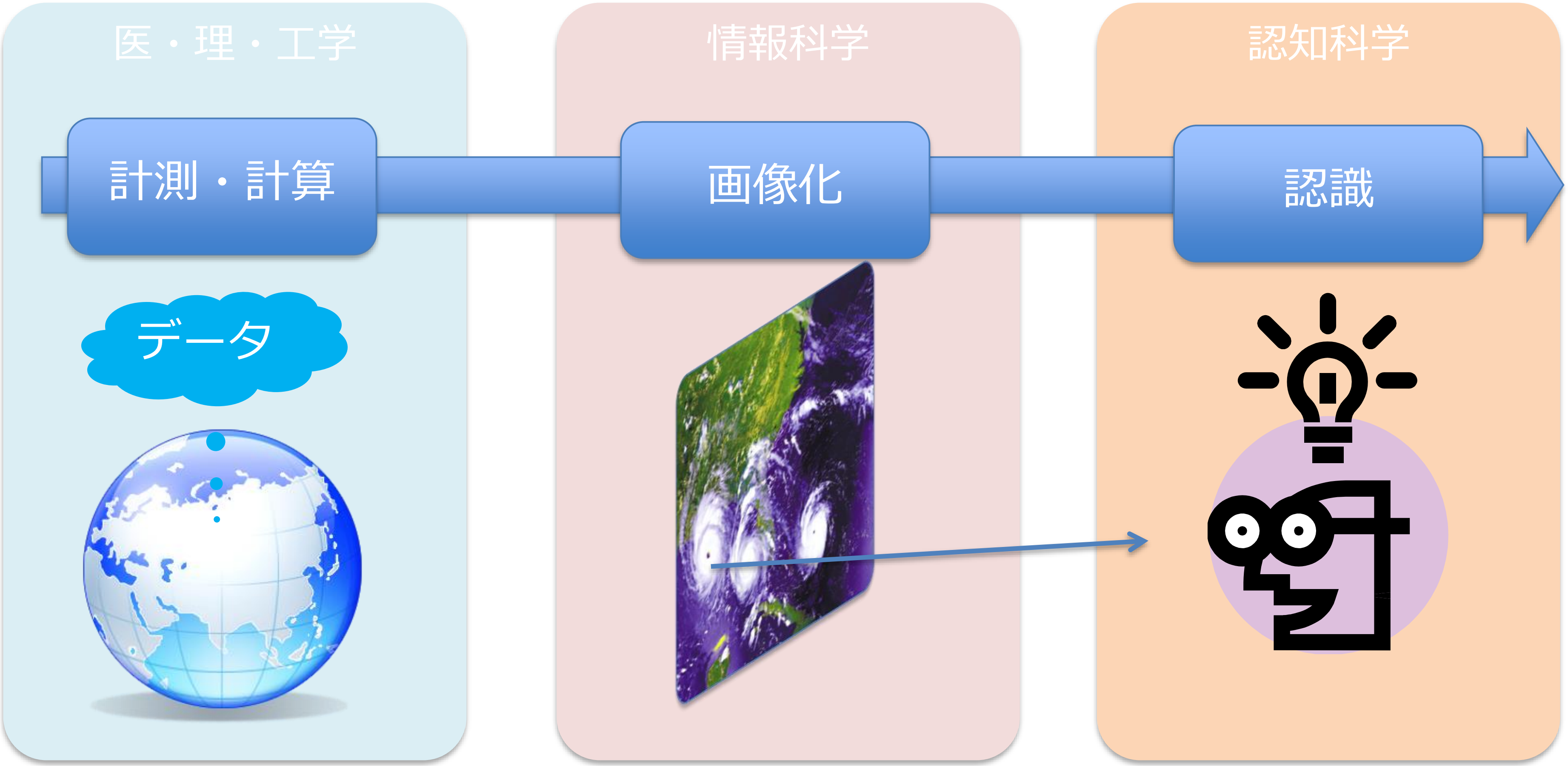


1. 融合可視化
2. 因果ボリュームデータの可視化
3. 興味領域の特定
4. 特定された領域での因果関係発見

可視化：ビッグデータ時代の科学を拓く

可視化の性能評価

可視化研究と性能評価

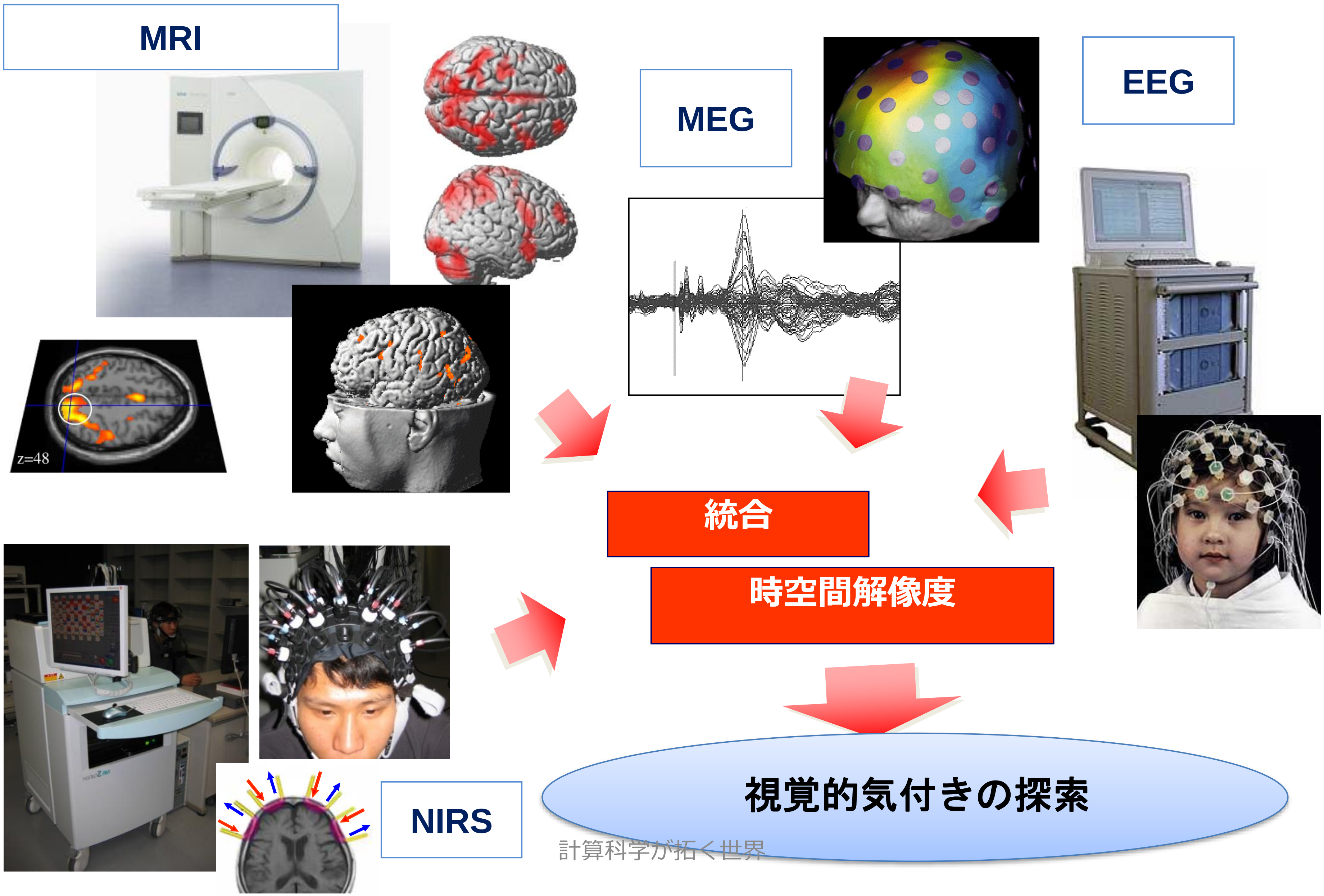


どんな現象をデータ化できたか？

どれほど効率よく画像化できたか？

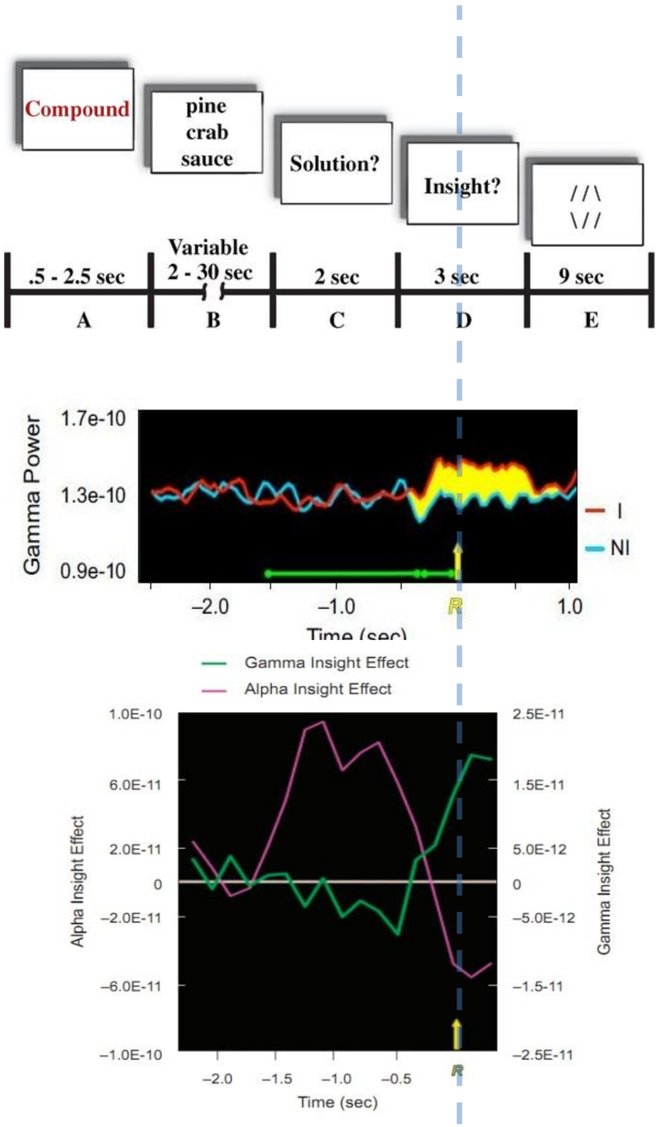
どれほどの気付きを得たか、どんな行動変容に結び付いたか？

脳機能計測装置による気付きの測定



アハ体験についての研究

Kounios, J, Beeman, M, “The Aha! Moment-The Cognitive Neuroscience of Insight,” 2009

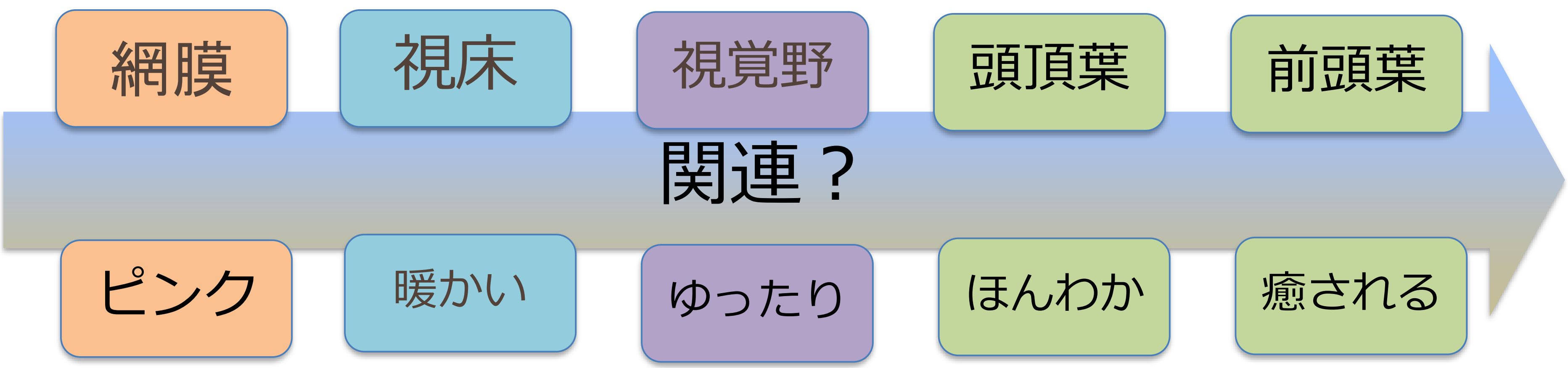


複合的遠隔性連想問題によるアハ体験

- ガンマ波 (約　40 ヘルツ) EEG 信号
- アルファ波 (約　10 ヘルツ) EEG 信号
- ひらめき時のガンマ波の組織的分布
- ひらめき時の機能的MRI分布

まとめ

- 科学的方法において重要な役割を果たす可視化研究
  - 粒子を用いた超拡張性可視化
  - 可視化による因果関係気付きの支援
  - 脳科学と連携する認知構造可視化



演習

- 指定されたTEDコンテンツを視聴し、「科学に問えるが、科学が答えられないもの」について回答せよ。
  - 以下サイトにアクセスして、200字程度で記述せよ

https://goo.gl/6BgPxh