人工知能とデータサイエンティストの役回り

樋口知之　（情報・システム研究機構　統計数理研究所）

大学共同利用機関法人と大学共同利用機関

文部科学省の国立研究所

全国に17設置

Universities/College

国立大学
83

私立大学
661

私立短大
468

Inter-University Research Institutes

自然科学研究機構 （国立天文台、…）

高エネルギー加速器研究機構

人間文化研究機構

情報・システム研究機構

統計数理研究所 (ISM)

国立情報学研究所 (NII)

国立遺伝学研究所

国立極地研究所

北川

樋口

喜連川

大学セクター

Bottom Up

甘利

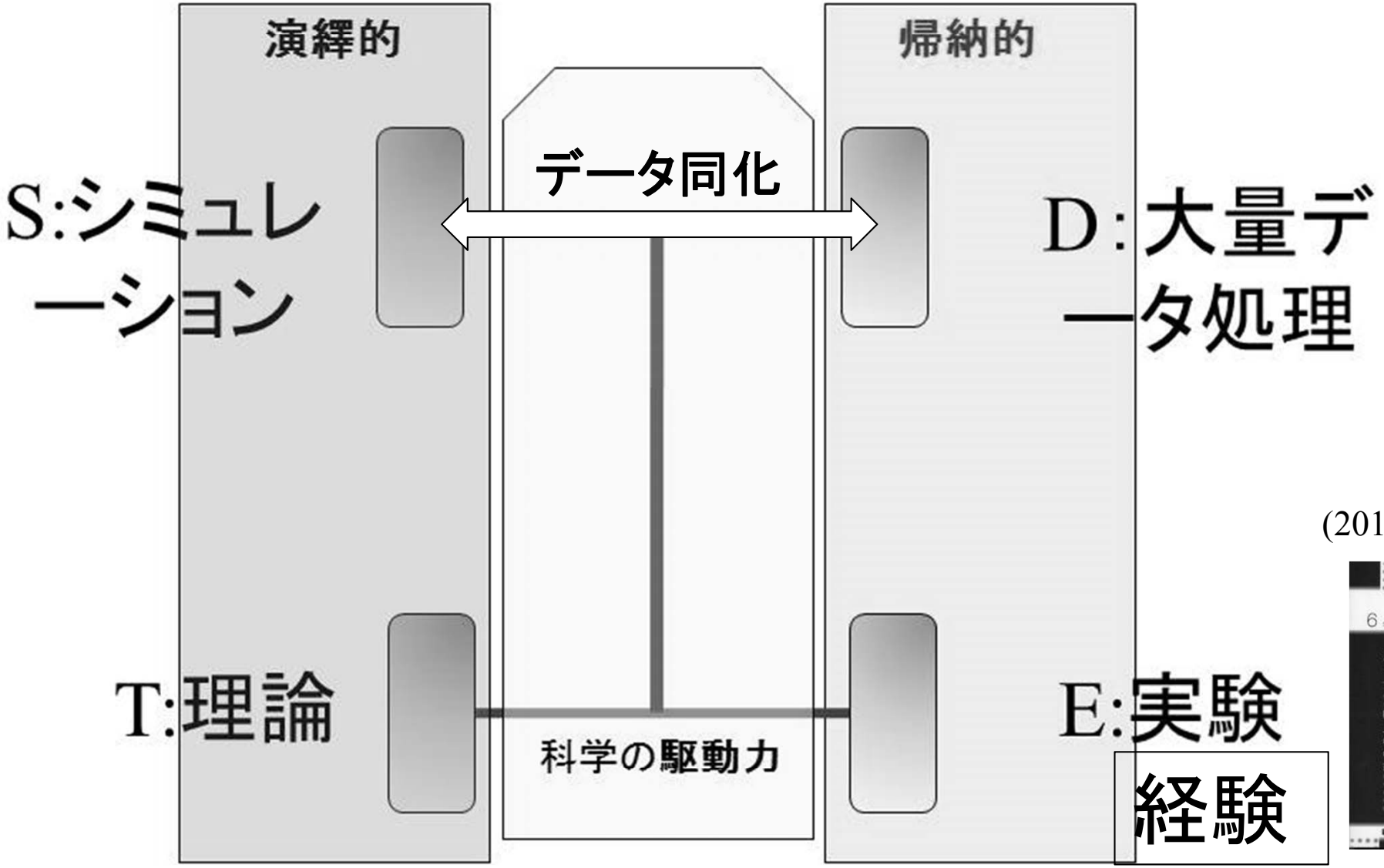
国立研究開発法人Top Down

理研、JAMSTEC, NIMS

アウトライン

1. ビッグデータと機械学習
2. 帰納法（データサイエンス）の弱点
3. 人工知能研究の今昔
4. データサイエンティストの役割
5. 温故知新

つなぐ：データ同化



(2011年9月刊行)

今日のテーマ：異質・相反する思考法

質問

Q. 統計学と確率論の違いは？

確率論は、偶然現象に対して数学的なモデルを与え、解析する数学の一分野

確率論は、統計学を記述する際の言語や道具

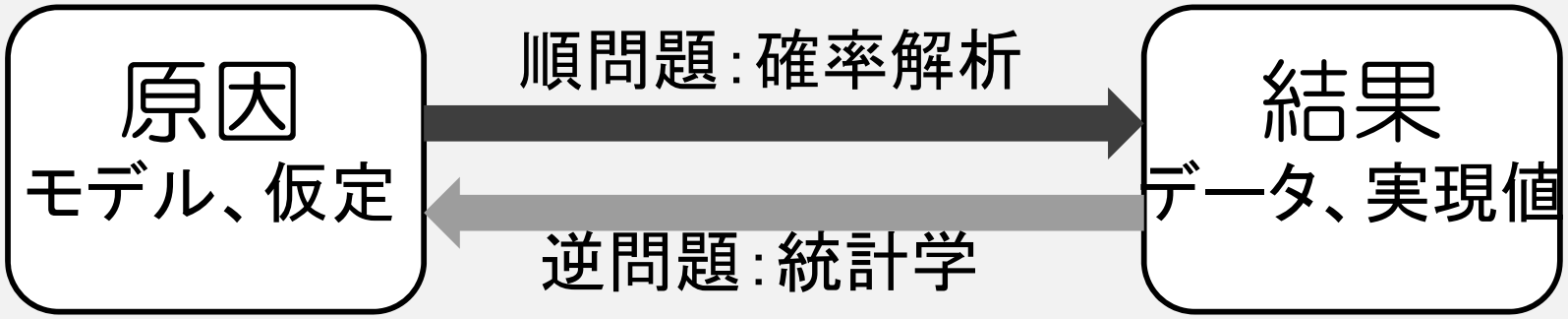
- 確率とは何か？
- 帰納法と科学的推論

統計学はメタ学問。人間くさく、深く、豊かな世界。

さいころ

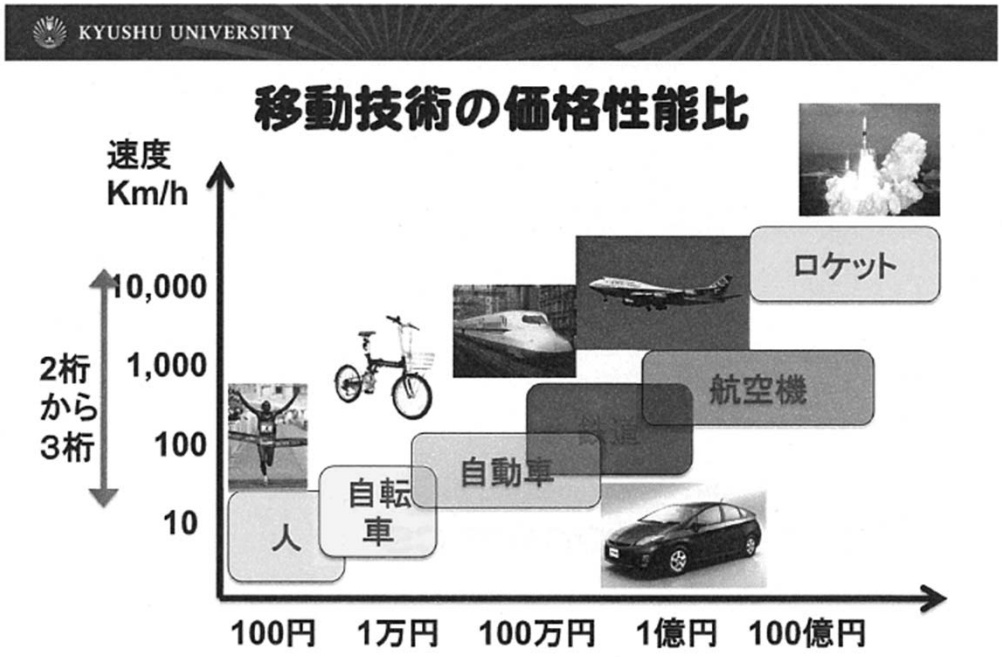
$$p(X=1)=\frac{1}{6}$$

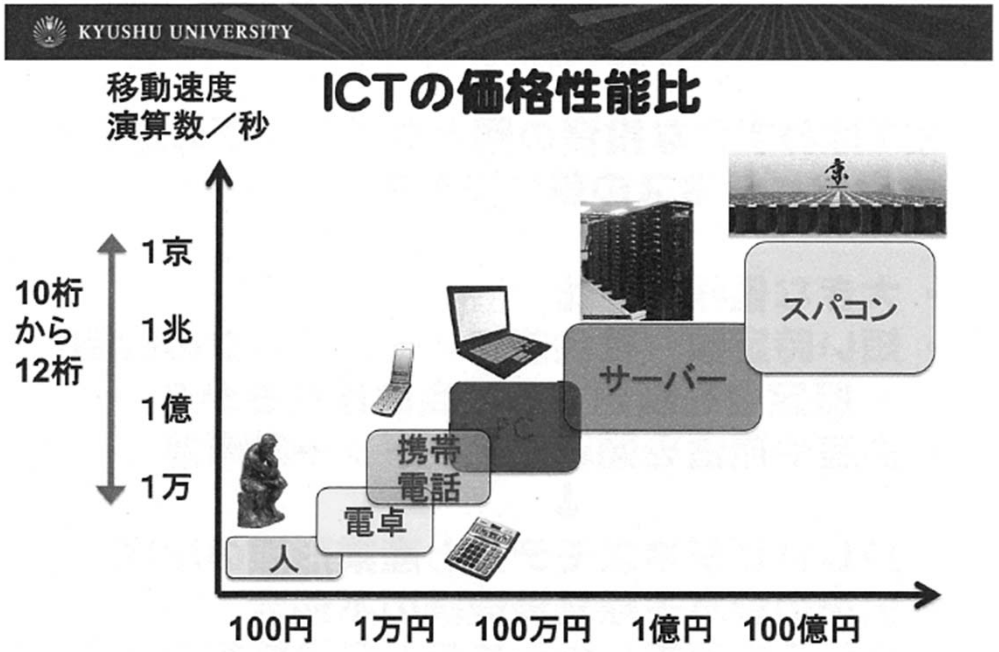
5, 4, 1, 3, 3,…



IT技術の破壊的浸透力

歴史に学べない時代
人類が経験したことのない時代

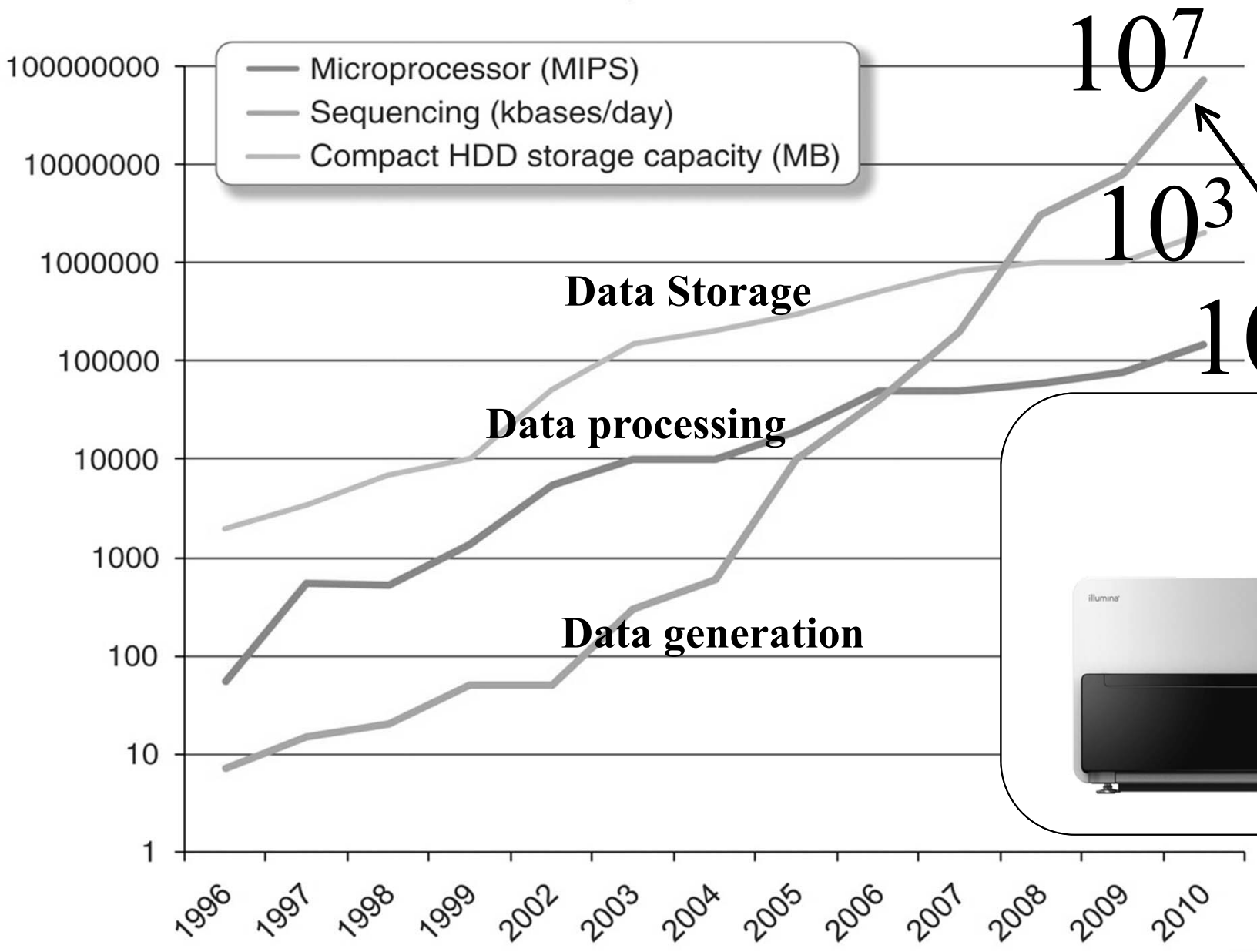




（九州大学理事・副学長　安浦寛人先生のスライドを借用・許可得）

Sequencing Progress vs Compute and Storage

Moore's and Kryder's Laws fall far behind



10^7

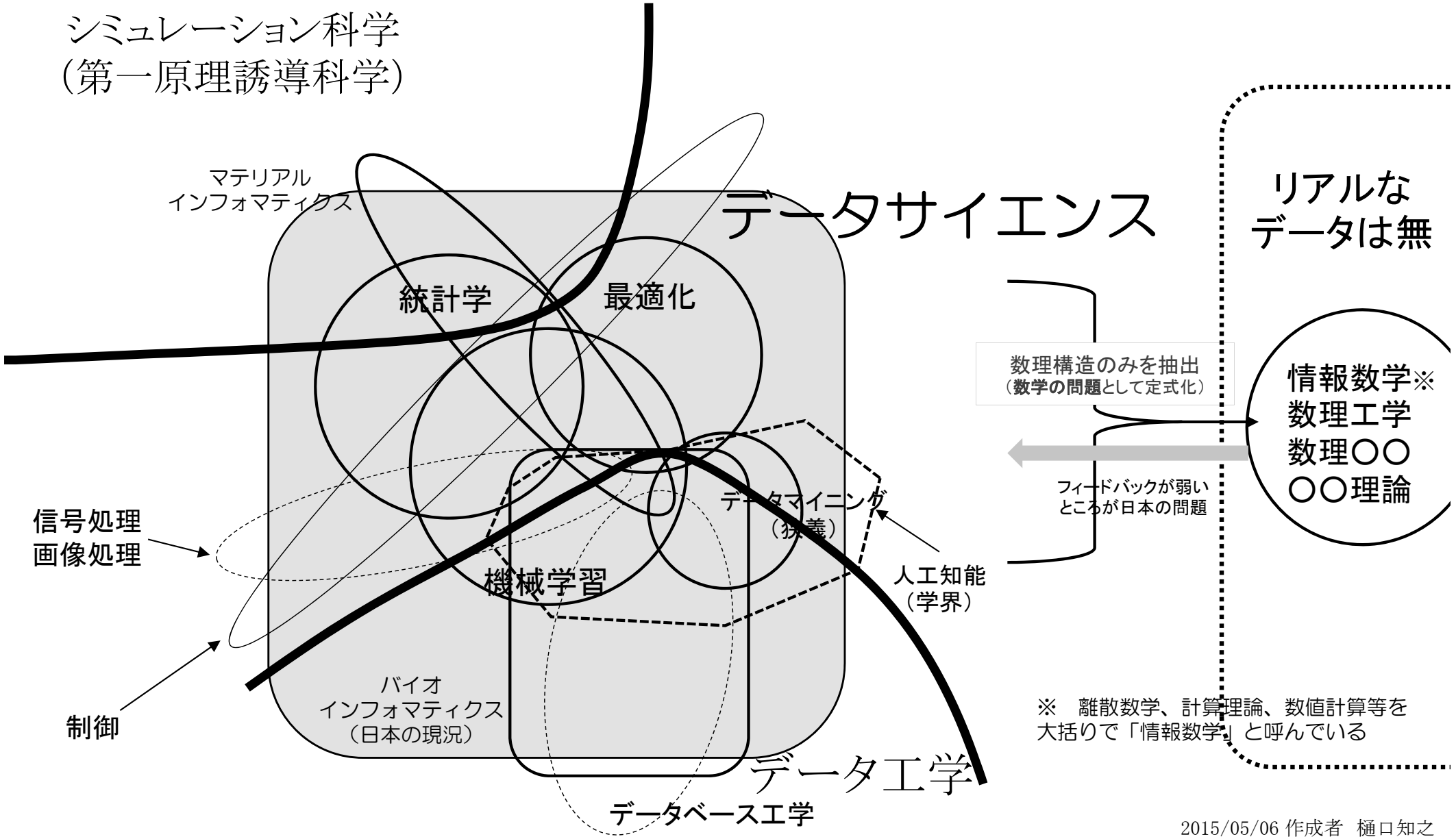
10^3

10^3

100Gb/day /sequencer

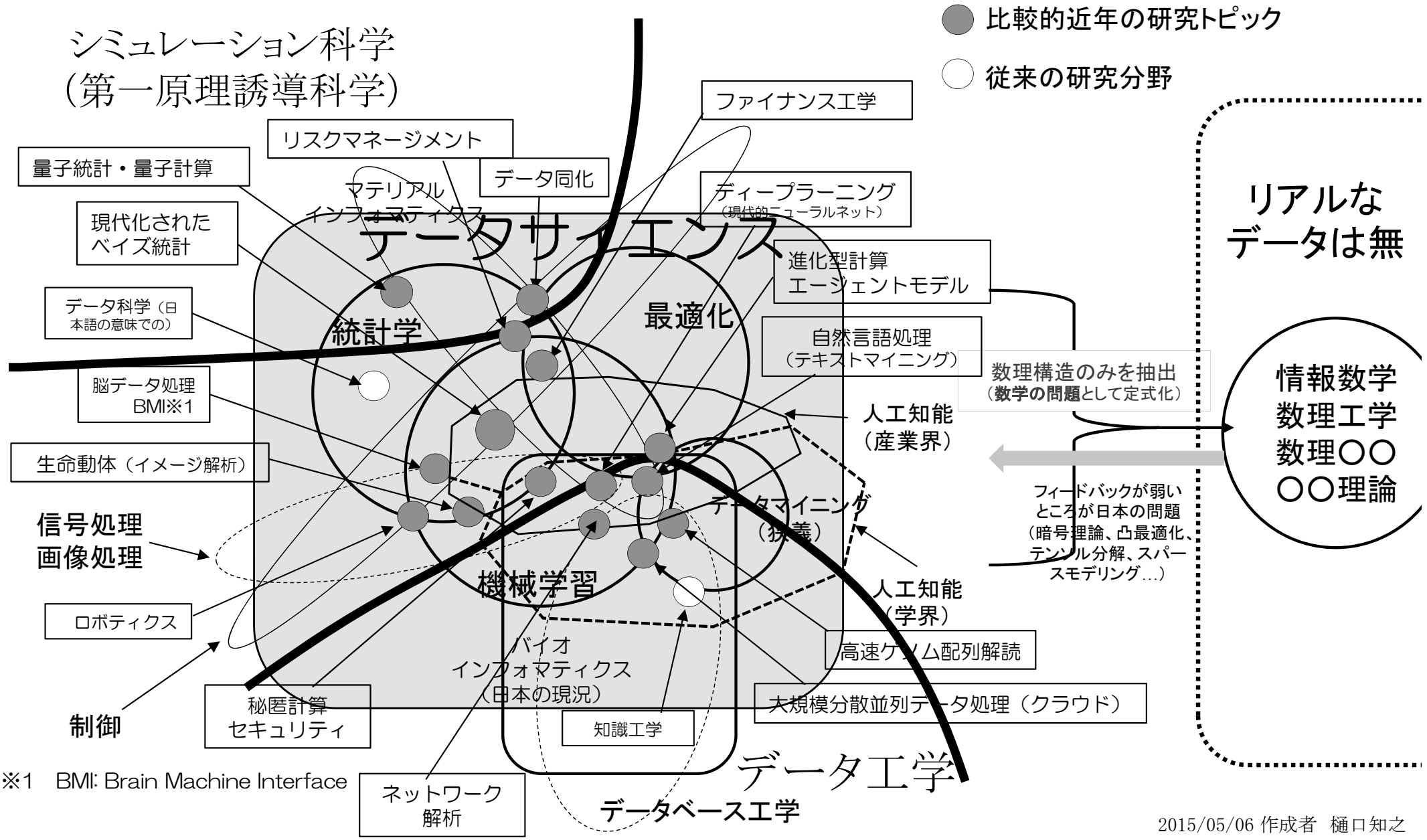
(Science, 2011)

データに関連した数理分野の俯瞰図



2015/05/06 作成者 樋口知之

データに関連した数理技術（研究トピック）のマップ



※1　BMI: Brain Machine Interface

2015/05/06 作成者　樋口知之

ボナンザ（保木 氏作）

2006年5月　第16回世界コンピュータ将棋選手権大会優勝

論理思考とデータ解析の組み合わせ

探索：巨大な状態空間の中の効率的、効果的な力づく探索

○全幅検索と選択検索のハイブリッド

機械学習：**6万局の棋譜データ**（※）から、評価関数のパラメータを自動生成。↑の“効率的”の源泉。

１億個のパラメータの最適化

※プロの公式戦の対局3万局と将棋倶楽部24の3万局

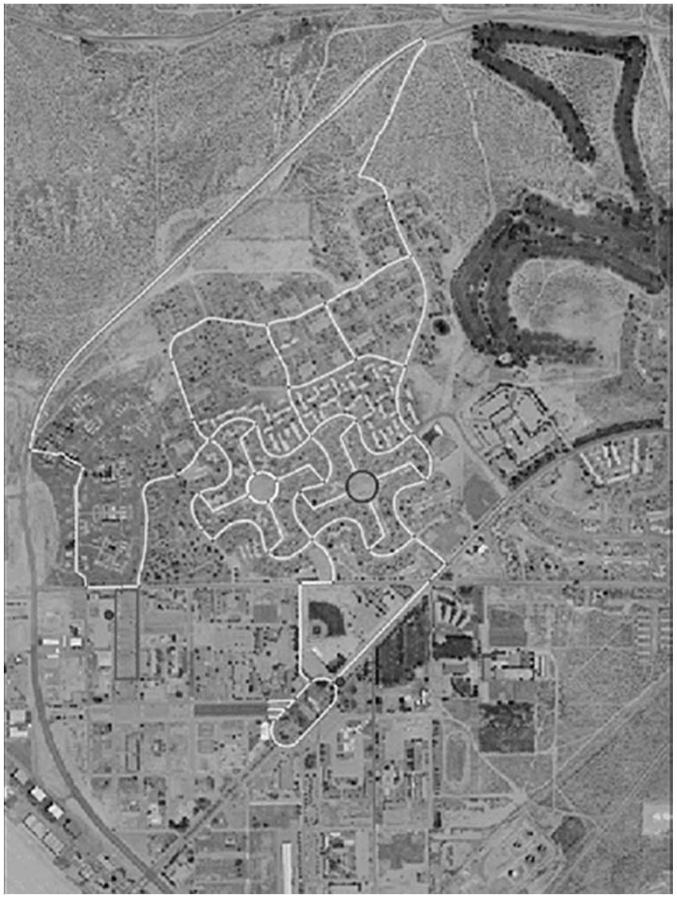
←ボナンザ囲いの発見

DARPA (米防衛高等研究計画局) *Urban Challenge*

完全ロボット操縦のカーレース

The DARPA Urban Challenge, which took place on November 3, 2007

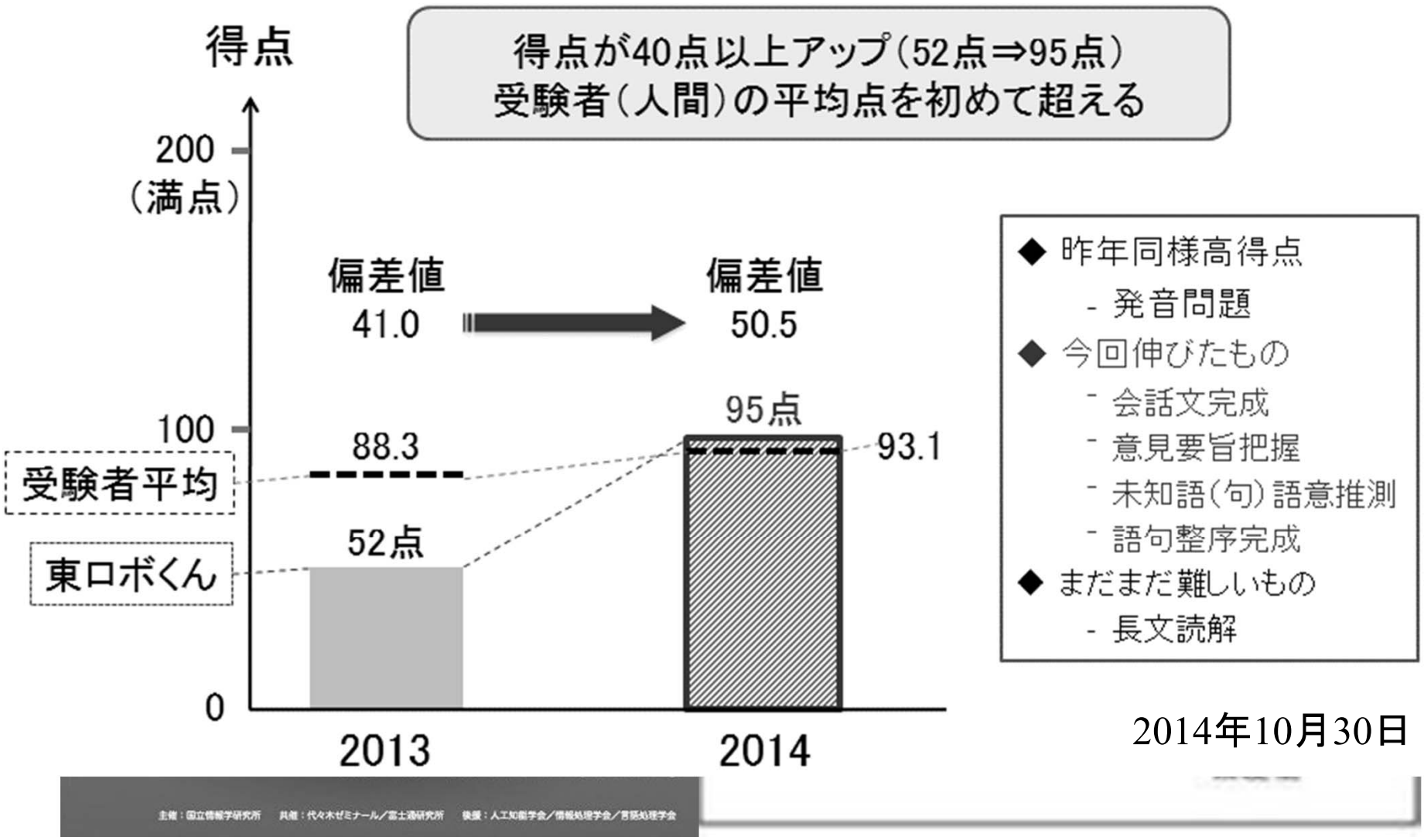
http://archive.darpa.mil/grandchallenge/



Google’s Driverless (Self-driving) Car

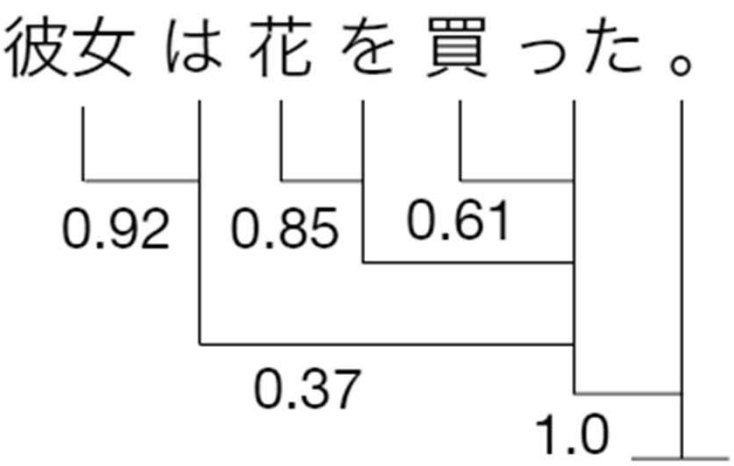
http://www.ted.com/talks/sebastian_thrun_google_s_driverless_car.html

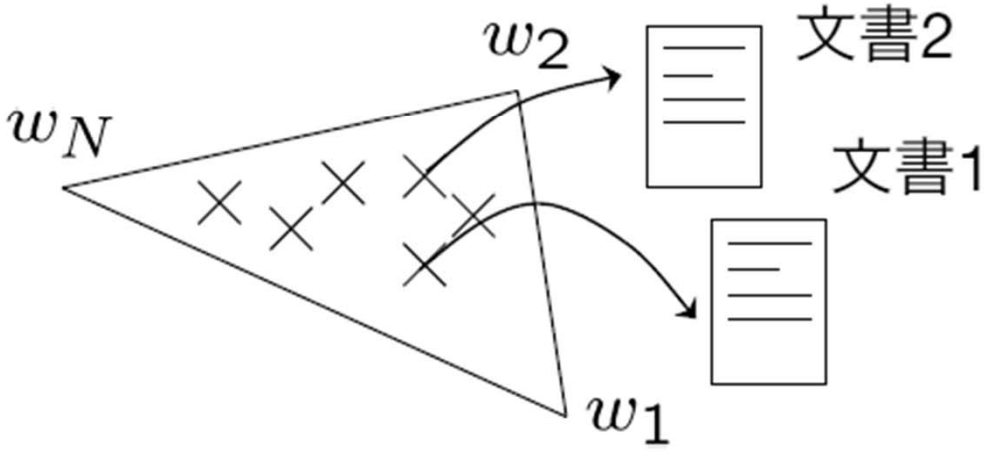
東ロボ君（新井教授@国立情報学研究所）



「自然言語処理」も統計的機械学習へ

- 「計算言語学」ともいわれる
  - 大量のテキストデータの統計的な分析に基づく
    - 形態素解析 (単語分割, 品詞付与)
    - 構文解析・係り受け解析
    - 統計的意味解析
    - 文書の統計モデルと情報検索 etc, etc …





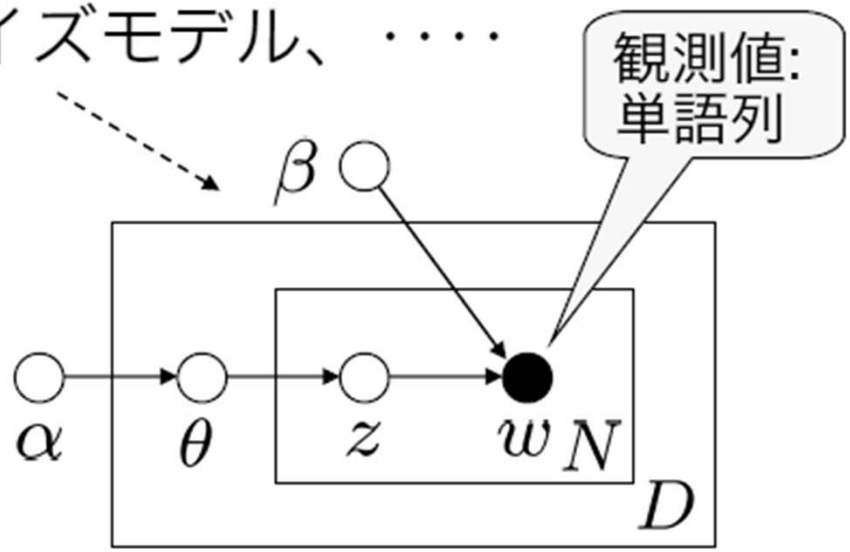
（統数研・持橋准教授のスライドから）

統計的自然言語処理

- 1990年代後半～からパラダイムシフト
  - 統計的機械学習の一部として重要な位置
- 論理式から、高度な統計モデルへ
  - チョムスキーの亡霊からの脱却
  - Webの登場と電子テキスト、計算資源の爆発的増大
  - 対数線形モデル、階層ベイズモデル、・・・・

$$p(t|\mathbf{x}, \mathbf{\Lambda}) = \frac{\exp\left(\sum_i \lambda_i f_i(\mathbf{x}, t)\right)}{\sum_{\mathbf{x}} \exp\left(\sum_i \lambda_i f_i(\mathbf{x}, t)\right)}$$

ある単語xの品詞
が形容詞である確率



これらの課題に共通する点

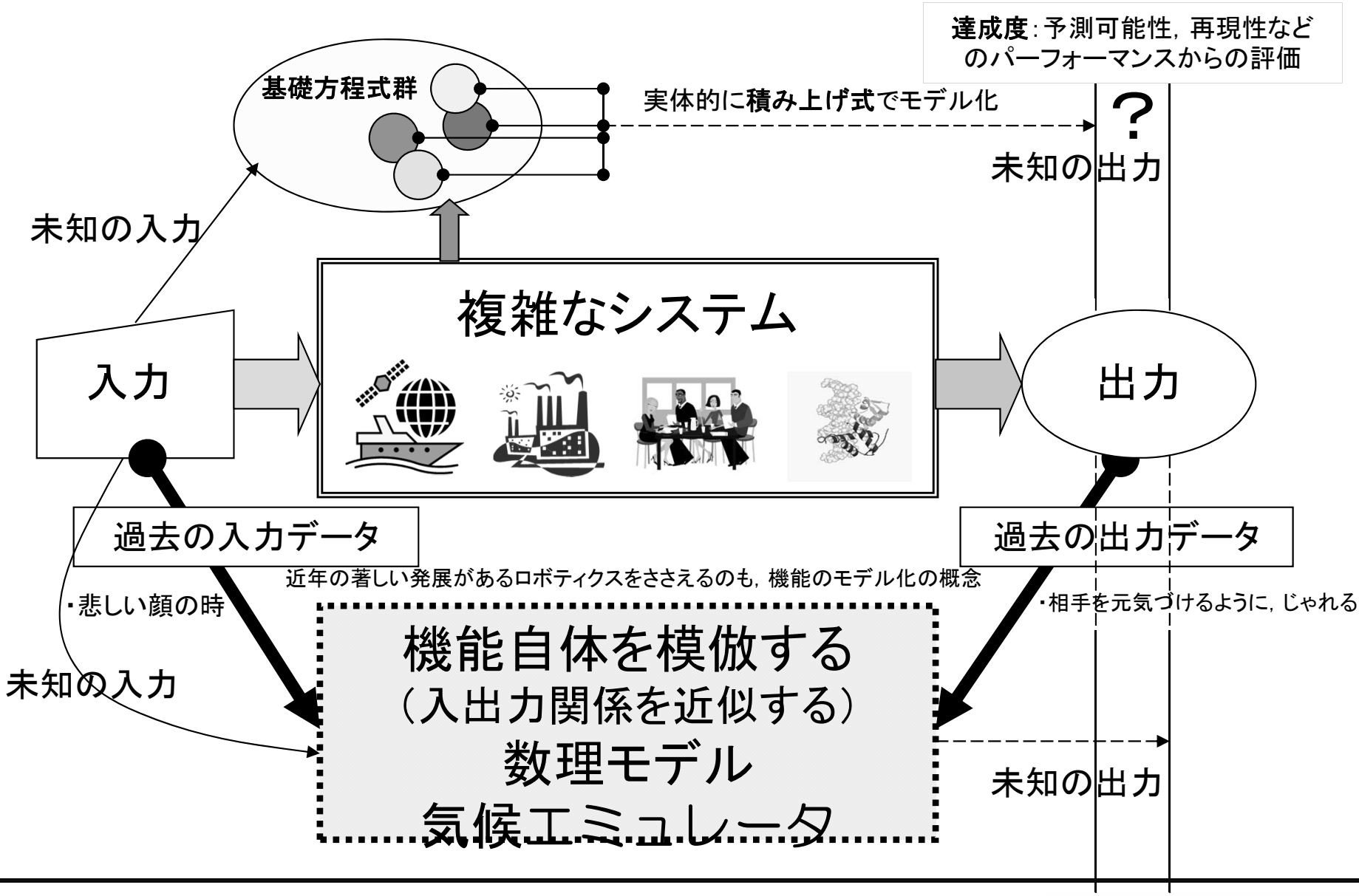
- タスクが明確：予測や判別
- 性能向上を通じた機能のモデル化
- ベストでなくベターを、それも早く求める

「認識科学」から「設計科学」へのシフト
「対象理解」から「機能の最適化」へ興味がシフト

- 対象に関する知識は常に不完全である。
- 現象の予測能力でもって研究の進め方を評価し修正する。
- 意志決定にはリスク解析（分析）をしっかりやらねばならない。

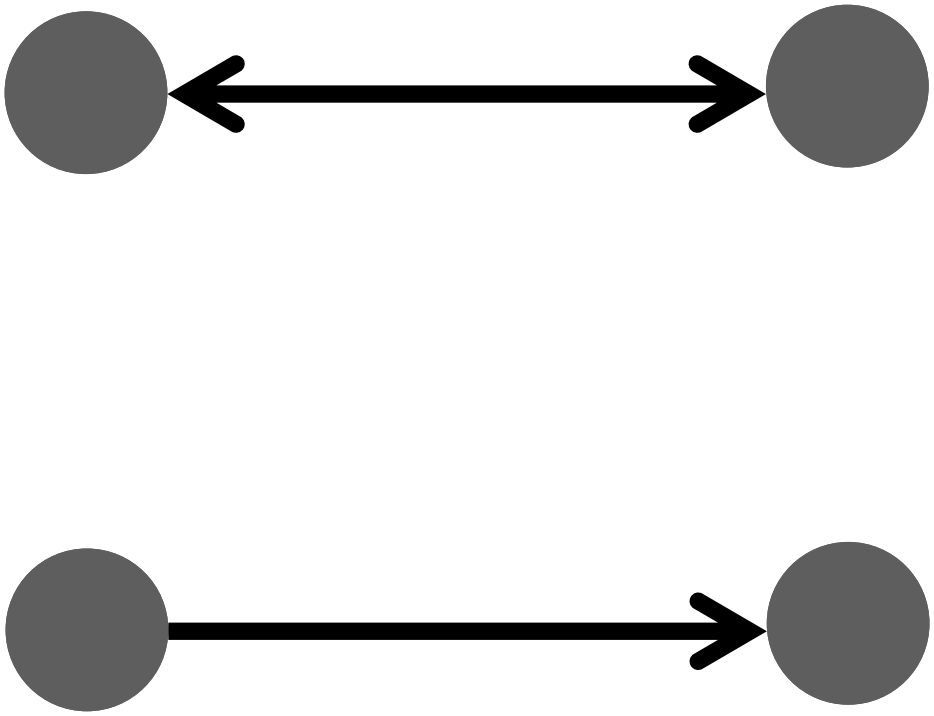
機能のモデル化：エミュレータ

高度情報社会におけるユニバーサルな研究課題の表現形



機械学習は万能か？
帰納法の弱点

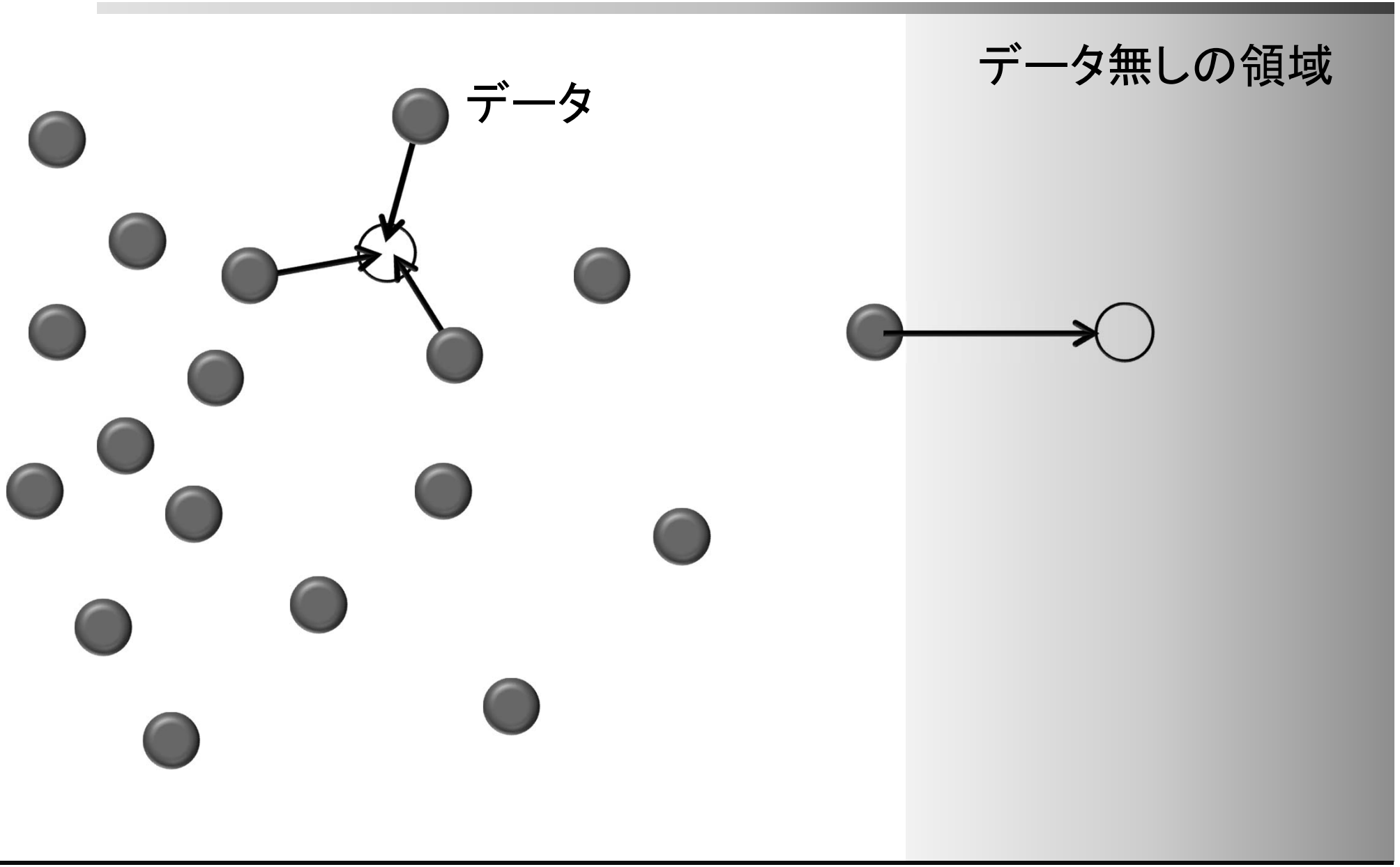
急所1：相関と因果



言っている内容は正しいが，
言い方は誤解を与える。

ビッグデータの時代には、暮らし方から世界との付き合い方まで問われることになる。特に顕著なのは、相関関係が単純になる結果、社会が因果関係を求めなくなる点だ。
「結論」さえわかれば、「理由」はいらないのである。

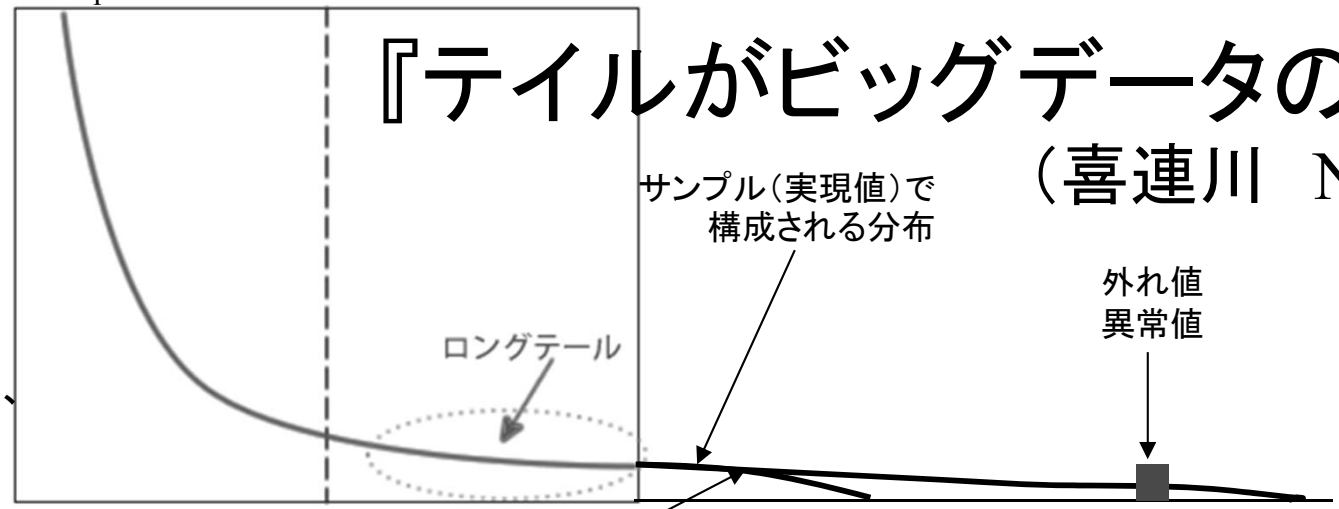
急所2：内挿と外挿問題



全てのデータを取り扱う意味

帰納法の弱点

http://sts-concrete.seesaa.net/index-2.html



実際は高次元なので、
端をみつけるのは
意外と大変

『テイルがビッグデータの醍醐味』
（喜連川　NII所長　談）

極値分布

河川の氾濫、最大風速、最大降雨量、
金融におけるリスク等の分布に適用

『端にこそイノベーションの卵』

- 新発見、ひらめき
- クレーム（PL法対応）　←エラー、故障、不正、侵入

そうでなければサンプリング（標本抽出）によって一部のデータを分析することで十分（費用対効果を最初から考えること）

機能のモデル化:“見よう見まね”を科学する

- “見よう見まね”のプロセスを加速する。

  ―体系化されていない研究分野において有効

- “見よう見まね”による完成の域がお手本を超える。

  ―贋作が“本物”を超える

日本人は伝統的に“贋作”を心から嫌う傾向が強いが、同じく漢字文化圏にある中国人で心ある人は、過去の文物に目を転じるとき、それがたとえ”贋作”であろうと判断できても、”本物”を超える出来栄えであるならば、自分の心と目を満足させるため、”本物”以上にその”贋作”を尊重して手に入れると伝えられている。

“贋作”と”本物”の、どちらが本物か実は分からない

第二の産業革命：知的労働の質がかわりつつある

佐野正博「技術の生存競争 ---「動力」に見る進化論」『週刊朝日百科 世界の歴史』第110号,朝日新聞社,1991,p.695

ワットの改良蒸気機関 （Wikipediaより）

深層学習：Deep Learning

２０～３０年周期のもりあがりには意味がある

世代交代に近い期間に勃興が同期する思想は要注意

データから価値を見いだす方法論の研究である、統計学を含むデータサイエンスはメタサイエンス。合理的に他者を納得させるツール。古くから、「統計学は科学の文法」と言われている。（ピアソン、1892）

ファクトの積み重ねで発展する自然科学の諸分野の中では、（コミュニティの）思考・価値観の投影が色濃くみられる特異な存在。→論文に、流行テーマが明らかに存在する： カーネル法、スパース学習

研究テーマの勃興（はやりすたり）の期間が、20～30年という、学術分野の世代交代の期間と近い場合は、同様の思想・価値観をもつ集団が（無意識に？＋計算機の発達に触発されて）再構成されたことが原因の可能性がある。

ベイズ統計の浮き沈み：実学、ノイズ、逆解析

1993年、樋口がベイズ研究の今後について（故）赤池先生に尋ねたところ
（赤池）「Neyman-Pearsonによれば、30年周期か何年周期で、Bayesの議論は復活してくる。彼自身がベイズ的なstructureを使ったんだから、最初。」

1763年: イギリスの牧師・数学者トーマス・ベイズ（1702 – 1761年）がベイズの定理を発見
18世紀後半から19世紀初頭: ラプラス
「天文学の誤差を含むデータをどう解釈するか」という極めて実務的な問題意識
1920年代： Fisher, Neyman, Pearsonらが数理統計の基礎を固める
冬の時代：
1）砲兵隊の誤差修正　2）電信自動接続の経路選択問題　3）保険料の算定方式
4）アラン・チューリングの独軍暗号解読　5）コルモゴロフと砲術　6）シャノンと暗号＋コミュニケーション
1950年代： Neo-Bayesian revival（Waldの統計的決定関数、von Neumann のゲーム理論の枠組みの影響）
1980年： 時系列に対するベイズモデリングで赤池先生が*Journal of Time Series Analysis*の第1巻第1号の巻頭を飾る
(Geman&Geman *IEEE*, 1984; Gelfand and Smith *JASA*, 1990)

参考文献

S.E. Fienberg, Bayesian Models and Methods in Public Policy and Government Settings, *Statistical Science*, 2011.
A.P. Dawid, Probability, Causality and the Empirical World: A Bayes–de Finetti–Popper–Borel Synthesis, *Statistical Science*, 2004.
同僚の**川崎准教授**にいろいろ教えてもらいました。
http://d.hatena.ne.jp/shorebird/20131228
シャロン・バーチュマグレイン著、異端の統計学 ベイズ, 草思社、2013.

魂を売る所業

日経新聞12月3日付サイエンス版「今どきの数学 （中）」

主観をまじえる手法に対し、

修正・洗練

能に育成

数学や科学の本流の立場からは「魂を売る所業だと批判されてきた」と樋口教授。

にもかかわらず支持されるのは役に立つから。ネット書店を

人工知能の隆盛（正確に言えば、『第三次ニューロブーム』）

■ この５年ぐらいの間に，人工知能分野は飛躍的な発展を見せている

- IBM Watson (質問応答システム) 2011年　人間のクイズ王に勝利.
  深層学習のパーツ無し。
  中身は、ビッグデータ解析結果の知識ベース＋エキスパートシステム
- 深層学習（Deep Learning）: 画像，音声など多くのコンテストでの圧倒的勝利 2011,2012～
  - 2011 交通標識の画像認識; (誤認識率 0.56%, ２位 1.16%)
  - 2012 ImageNet 一般物体認識（誤認識率 15.3%, ２位 26.2%）
    など

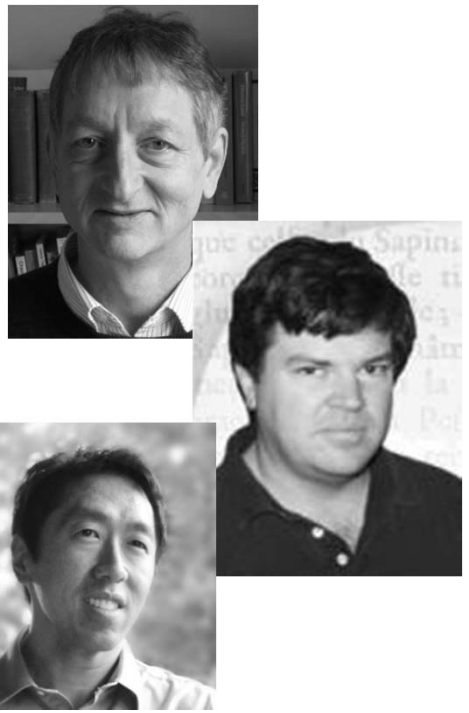
- 深層学習への注目度高まる. IT企業が研究に注力
  - Google: Google Brain Project
  - Facebook: Facebook AI Research
  - Baidu

（統数研・福水教授のスライドから改編）

コア技術：機械学習

■ 現在の人工知能の隆盛を支えるのは**統計的機械学習**の技術

不確実性（確率的構造）



- 深層学習は、機械学習コミュニティが現在の流行を作る
  中心的研究組織
  - Toronto大 Hintonグループ
    第2，第3ニューロブームを牽引
  - Facebook AI Research: Director, Yann LeCun
  - Google Brain Project
    Co-inventor, Andrew Ng; Jeff Hintonも参加
  - Baidu: Chief Scientist, Andrew Ng

  これらの研究者は，NIPS, ICMLなどの国際会議を
  中心とする機械学習コミュニティで活動

- 自然言語処理、画像・映像、音声にも、統計的機械学習で開発された技術が多く使われる（例：ノンパラメトリックベイズ）

（統数研・福水教授のスライドから改編・追記）

人工知能技術の現状と動向

■ 深層学習による認識問題における性能の飛躍
画像特徴の無教師学習、一般物体認識、文字・画像認識、音声認識
化合物反応予測

■ 現状の事例
- DeepFace (Facebook)：ある人の顔を，Facebook内の画像から自動検索，タグ付け
  人間レベルの認識性能
  (参考：　認識率97%, FBIのシステム85%)
- Googleなど：　画像の中の物体にタグ付け，説明
- 対話システム
- 音声認識
- 推薦システム(e.g. YouTube)
- 自動運転（Google)
- ヒューマノイドロボット（Google）
  http://www.foxnews.com/tech/2015/08/18/life-size-humanoid-robot-takes-walk-in-woods/

まだ、認識計算（Perceptual Computing）が中心

（統数研・福水教授のスライドから改編）

人工知能研究および研究開発プロジェクトの歴史

記号・論理

知識・統計

ニューロ

通産省ICOT（1982～1992）
（電総研）

通産省RWC
(1992～2001)

1998～2000　発見科学（有川）
2007～2009　情報大航海（喜連川）
2005～2010　情報爆発（喜連川）

1960年代　記号処理方式
記号処理のためのルールや数式をプログラム化し思考や推論など人間が行う情報処理を行わせる

1980年代　エキスパートシステム
専門家の知識やノウハウをルール化し、コンピュータに処理を行わせる

第5世代

2000年代　統計的アプローチ
膨大なデータをベイズ理論に基づく統計的手法で計算し自らルール生成し情報処理する

$$P(A|H)=\frac{P(H|A)P(A)}{P(H)}$$

2007年出版

統計的機械学習

AI冬の時代

1998年Google創業

2004年Facebook創業

1950　1960　1970　1980　1990　2000　2010

AI冬の時代

1958年代　パーセプトロン
脳の神経活動を数式モデル化しコンピュータに処理させる初歩的なニューラル・ネットワーク

1980-90年代　第2次ニューロブーム
ニューロ・ファジー家電製品

2006年　ディープラーニング
脳科学の研究成果を取り入れより忠実に脳の神経活動を再現

＊http://blogs.itmedia.co.jp/itsolutionjuku/2015/07/post_105.html　を参考として作成
＊統数研・福水教授のスライド内容を一部借用

過去を再考（1）

- 80年代：（古典的)AIの展開
  エキスパートシステム，第5世代，高速な推論エンジン
  本質的壁： フレーム問題，知識ベースの不足

- 80年代中～90年代中： 第2次ニューロブーム．多層パーセプトロン
  データに基づく非線形モデルによる推定
  パターン認識などで高い性能
  計算に莫大な時間
  試行錯誤による構造，パラメータチューニング
  理論構築困難（精度保証など）

- 90年代中～現在： 統計的機械学習の隆盛
  確率モデル＋計算アルゴリズムに基づく推論， 理論的裏付け

過去を再考(2)＋今後のAI研究に重要な視点

○2000年代後半から： 深層学習，ニューラルネットの復権
多くの層を使うことにより，大量の訓練データで学習可能

- 第2次ブームと何が違うか？
  - 学習アルゴリズムはあまり変わらない
  - 計算機の高速化，特にGPUにより安価に並列化可能
  - 大量のデータ： ITの発達により多くのデータが電子化されて利用できる

ビッグデータの登場と工夫により過学習を避けることに成功

- 「黒魔術」
  - 試行錯誤による，構造，パラメータチューニング
  - 理論解析／精度保証が困難

クラウドビジネスとSNSの浸透

これらは変わっていない．

Pendulum swings

●ノウハウの泥臭い集合体だと飛躍につながらない
⬇
●次を見据えて、技術を体系化し俯瞰するフェーズ
●見通しよく技術開発するセンスと地力
⬇
●確率モデルとアルゴリズムの潜在力

○数理基盤に根差した方法の必要性
今後10年の重要技術： 機械学習の中でも特に，
「確率モデル＋最適化、サンプリング（モンテカルロ計算）」による高性能な推論

（統数研・福水教授のスライドから改編・追記）

工夫の一例

- プレトレーニング
- ドロップアウト
- RNN (Recurrent NN, Recursive NN)
- AutoEncoder
- Convolutional NN　(Pooling)

データサイエンスの王道芸・宿命からは逃げられない

- 次元圧縮 （次元削減）
- 効果的非線形変換
- 超平面探索
- 近傍データを利用した内挿操作
- フィルタリング
- 残差の利用 (AR→ARMA）

そもそもできないものは絶対できない！

IEEE Interview: MJ (UCB) 58才

ディープラーニングはどうなのか？

その後、ベイズ統計の隆盛はとどまらない。何が解放したのか？

- 優れた、計算機集約的なアルゴリズムの提案 (MCMC, PF,ノンパラベイズ)
- 計算機の発達（1990年以降はスパコン）
- 機械学習コミュニティとの協働
- 自然科学・工学だけにとどまらない応用範囲の拡大とおもしろい実問題の探索

- ビッグデータ
- ハイエンドな計算機環境のコモディティ化：GPGPU とクラウド

少なくとも、定型のデータに対する広義のPerceptional Computing ＋教科学習においては主流になる可能性

特徴選択の作業が軽減

ビッグデータの登場により避けられない、けっこう地味なタスク（前処理）：

データクレンジング（異常値・欠損値処理）、データエディティング、データキュレーション

帰納法だけでインサイト・マシンはつくれるのか？

DSに求められるスキルセット

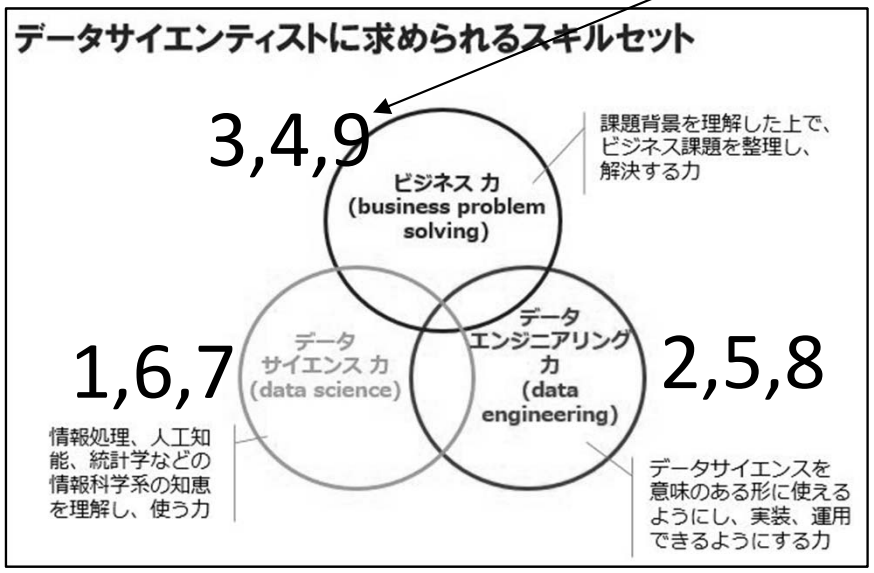
日本学術会議情報学委員会
E-サイエンス・データ中心科学分科会提言
「ビッグデータ時代の人材育成」（平成26年9月11日）

【ビッグデータ活用に必要な3大要素技術】

1. ビッグデータ処理技術
2. データ可視化
3. データ解析法

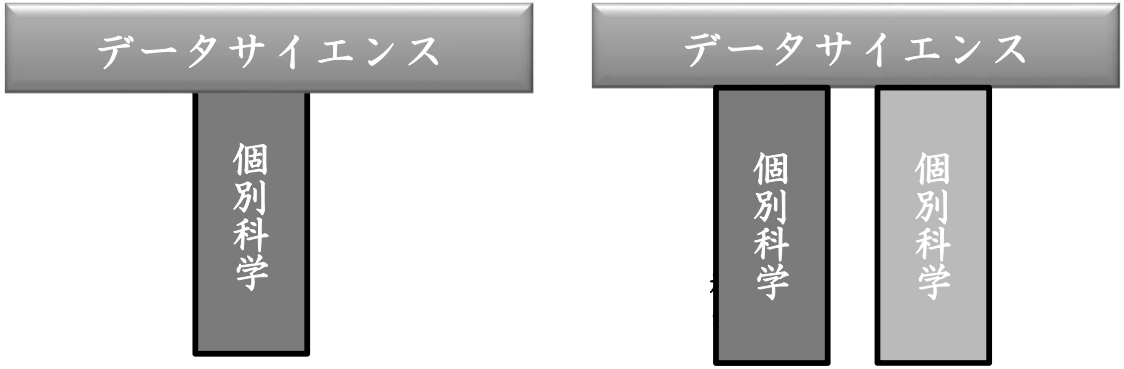
【データサイエンティストの要件＝データリテラシー】

1. ビッグデータ活用に必要な３大要素技術の習熟
2. セキュリティの知識習熟と研究
3. 研究倫理の徹底
4. 戦略立案能力，問題発掘・企画能力，問題解決能力
5. データ取集能力
6. データの裏にある真実を見抜き関連するデータを見出す能力
7. キュレーション能力
8. データ分析結果の業務や事業への実装能力
9. 異分野研究者・事業者との連携能力



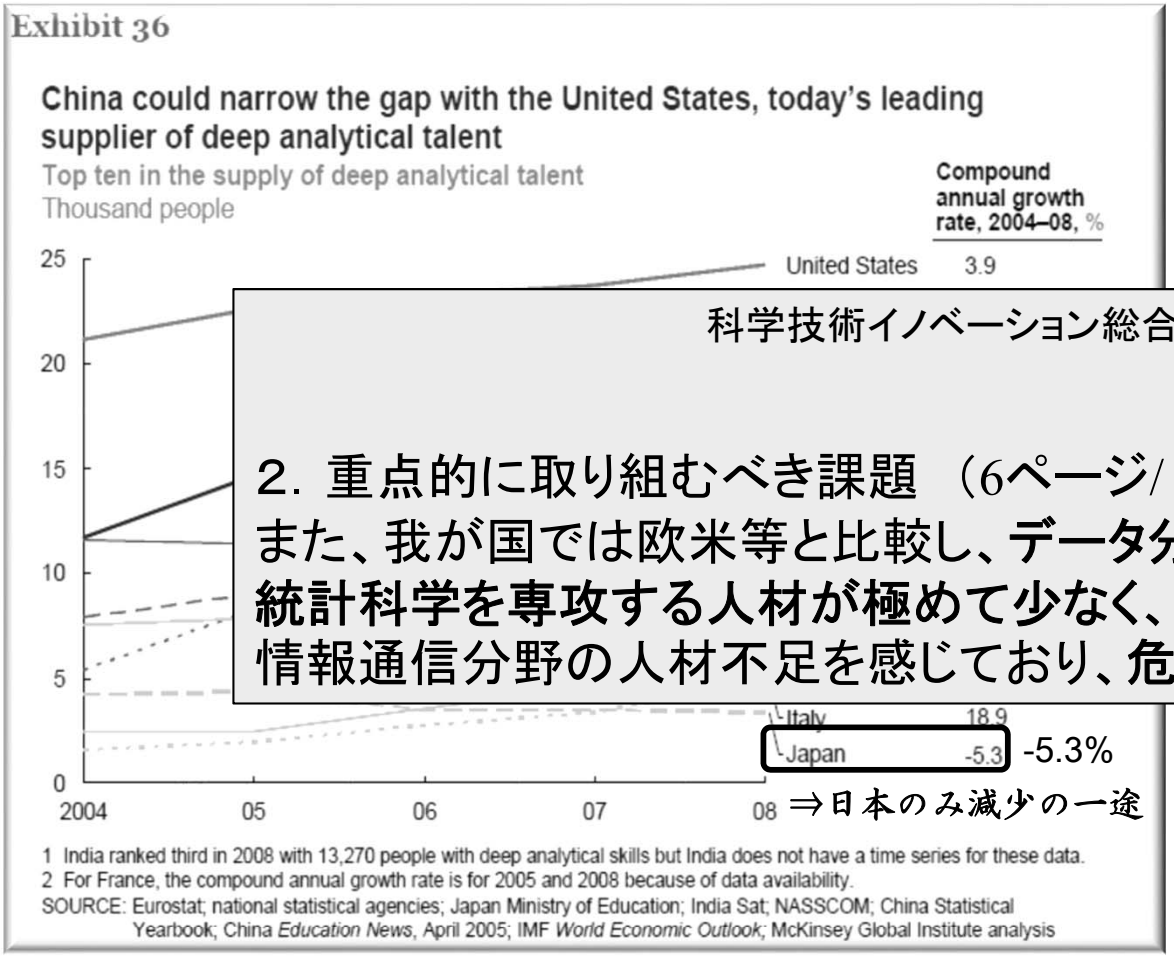
データサイエンティスト協会が定めたスキルセット（2014年12月）

T型人材，π型人材の育成が必須



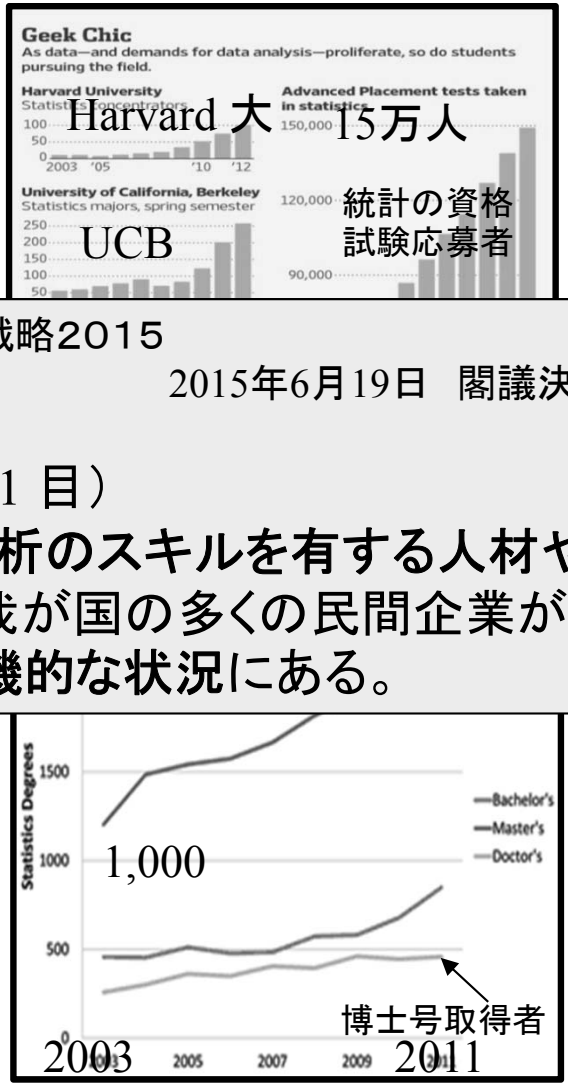
膨れ上がるデータサイエンティストへの期待

米国との差を縮めつつある中国



MGI（McKinsey Global Institute)リポートより

米国における統計学の推移



←米国において高まる学生の統計学の関心
(Wall Street Journal.)

米国における統計学の学士・修士・博士号取得者数の推移
(2003-2011)
http://magazine.amstat.org/blog/2013/05/01/stats-degrees/

科学技術イノベーション総合戦略2015
2015年6月19日　閣議決定

2．重点的に取り組むべき課題　（6ページ/ 81 目）
また、我が国では欧米等と比較し、**データ分析のスキルを有する人材や統計科学を専攻する人材が極めて少なく**、我が国の多くの民間企業が情報通信分野の人材不足を感じており、**危機的な状況**にある。

	Physicians and Surgeons	Physicians and surgeons diagnose and treat injuries or illnesses. Physicians examine patients; take medical histories; prescribe medications; and order, perform, and interpret diagnostic tests. They counsel patients on diet, hygiene, and preventive healthcare. Surgeons operate on patients to treat injuries, such as broken bones; diseases, such as cancerous tumors; and deformities, such as cleft palates.	Doctoral or professional degree	This wage is equal to or greater than $187,200 per year.

	OCCUPATION	JOB SUMMARY	ENTRY-LEVEL EDUCATION	2012 MEDIAN PAY
	Actuaries	Actuaries analyze the financial costs of risk and uncertainty. They use mathematics, statistics, and financial theory to assess the risk that an event will occur and they help businesses and clients develop policies that minimize the cost of that risk. Actuaries' work is essential to the insurance industry.	Bachelor's degree	$93,680
	Mathematicians	Mathematicians use advanced mathematics to develop and understand mathematical principles, analyze data, and solve real-world problems.	Master's degree	$101,360
	Operations Research Analysts	Operations research analysts use advanced mathematical and analytical methods to help organizations investigate complex issues, identify and solve problems, and make better decisions.	Bachelor's degree	$72,100
	Statisticians	Statisticians use statistical methods to collect and analyze data and help solve real-world problems in business, engineering, the sciences, or other fields.	Master's degree	$75,560

Publish Date: Wednesday, January 8, 2014

DSのスキルレベルと育成人数/年の目標

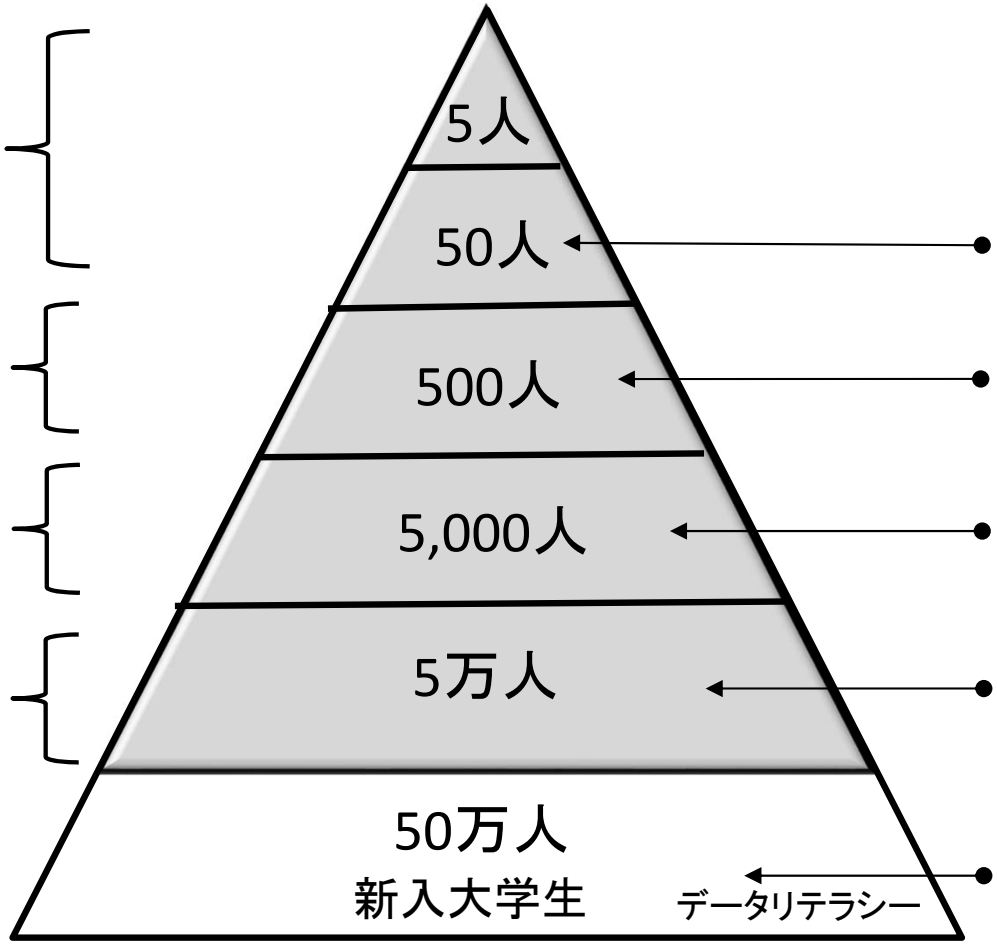
データサイエンティスト協会が定めたスキルレベル（2014年12月）

1．業界を代表するレベル
Senior Data Scientist

2．棟梁レベル
(full) Data Scientist

3．独り立ちレベル
Associate Data Scientist

4．見習いレベル
Assistant Data Scientist



統計数理研究所の将来の公開講座レベル（イメージ）

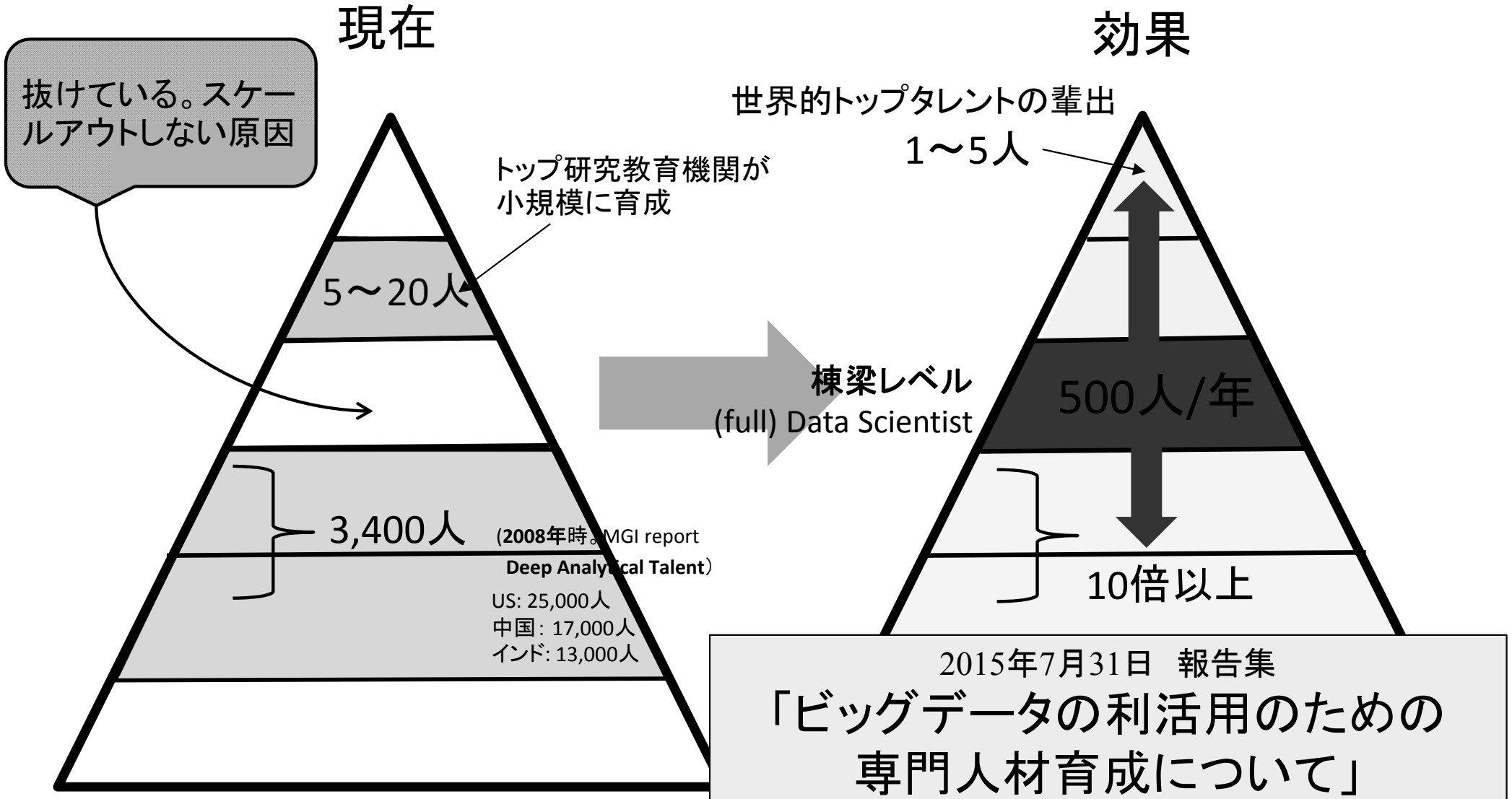
超上級 — RSS/JSS
上級 — 1級
中級 — 準1級
初級 — 2級
入門 — 3級

現在開講の講座レベル

統計検定のレベルとの対応

棟梁レベルDSを組織的・集中的に育成

DS育成における根源的問題点の解決



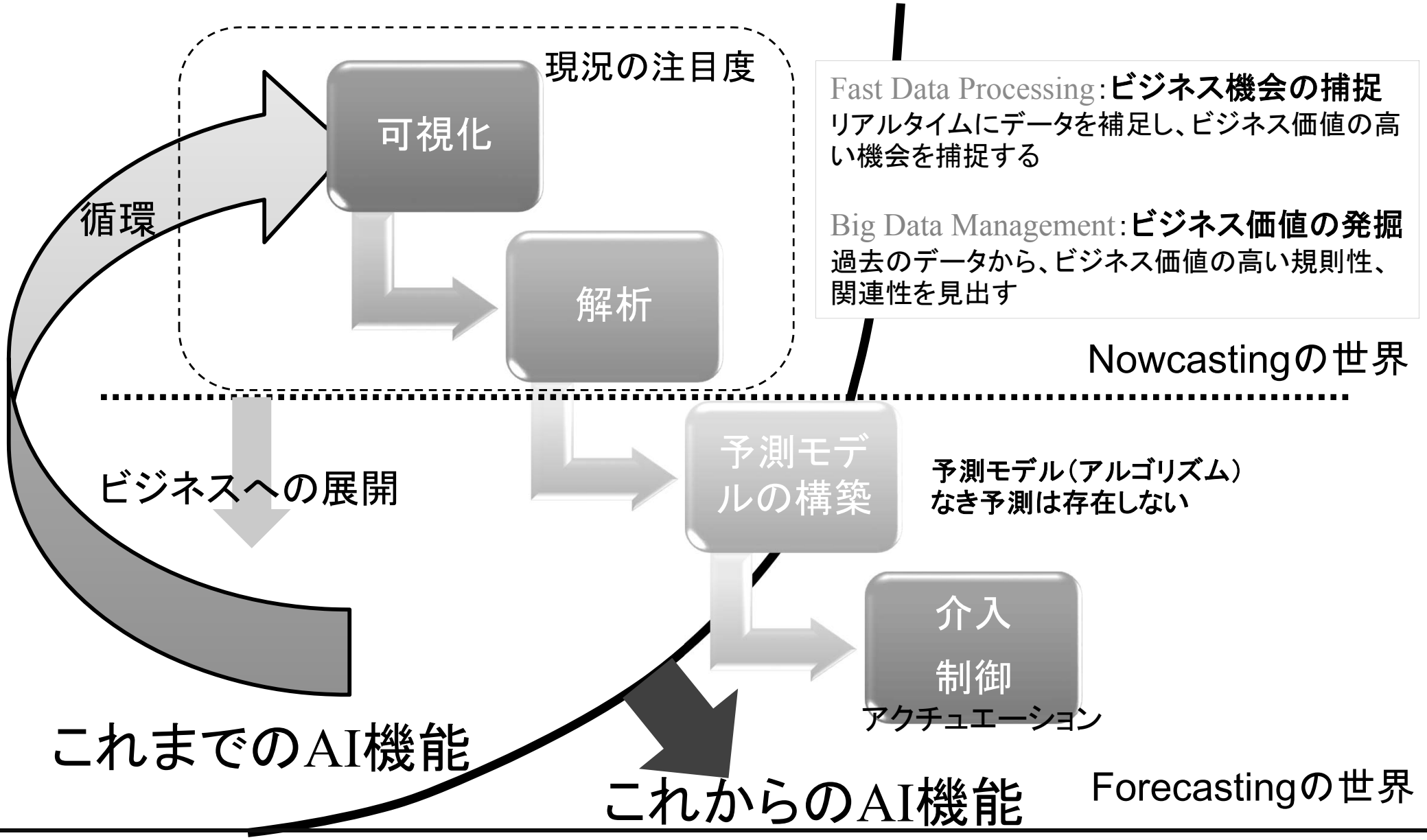
2015年7月31日　報告集

「ビッグデータの利活用のための専門人材育成について」

情報・システム研究機構

ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会

ビッグデータ利活用の４つのステージとＡＩの機能



個人への直接的な関与：Personalization

Achieve technologies that personalize products and services

personal, unique, individual, characteristic

個人、個性、個別、固有

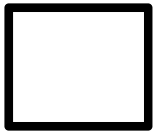
Conditioning

$$p(\mathbf{y} \mid \theta_1, \ldots, \theta_P)$$

Personal Services

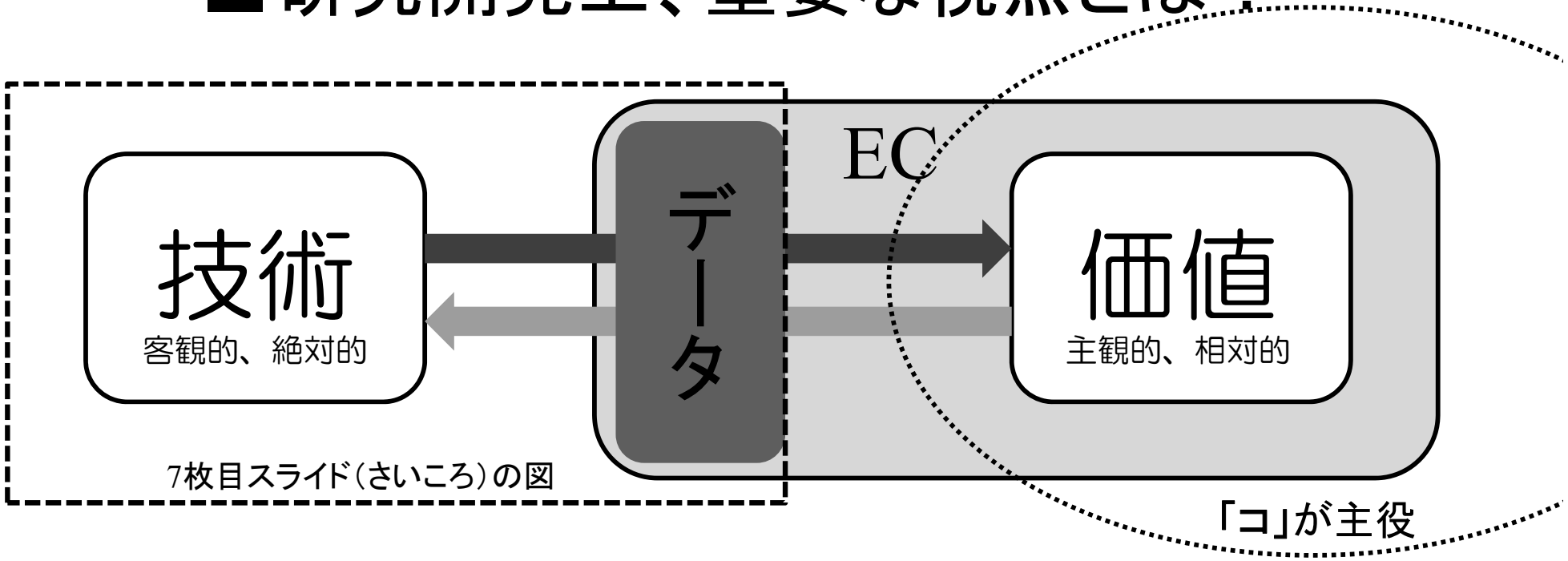
Real-Time Bidding

Personal Medical **Services**

cookie information 

帰納と演繹

■研究開発上、重要な視点とは？

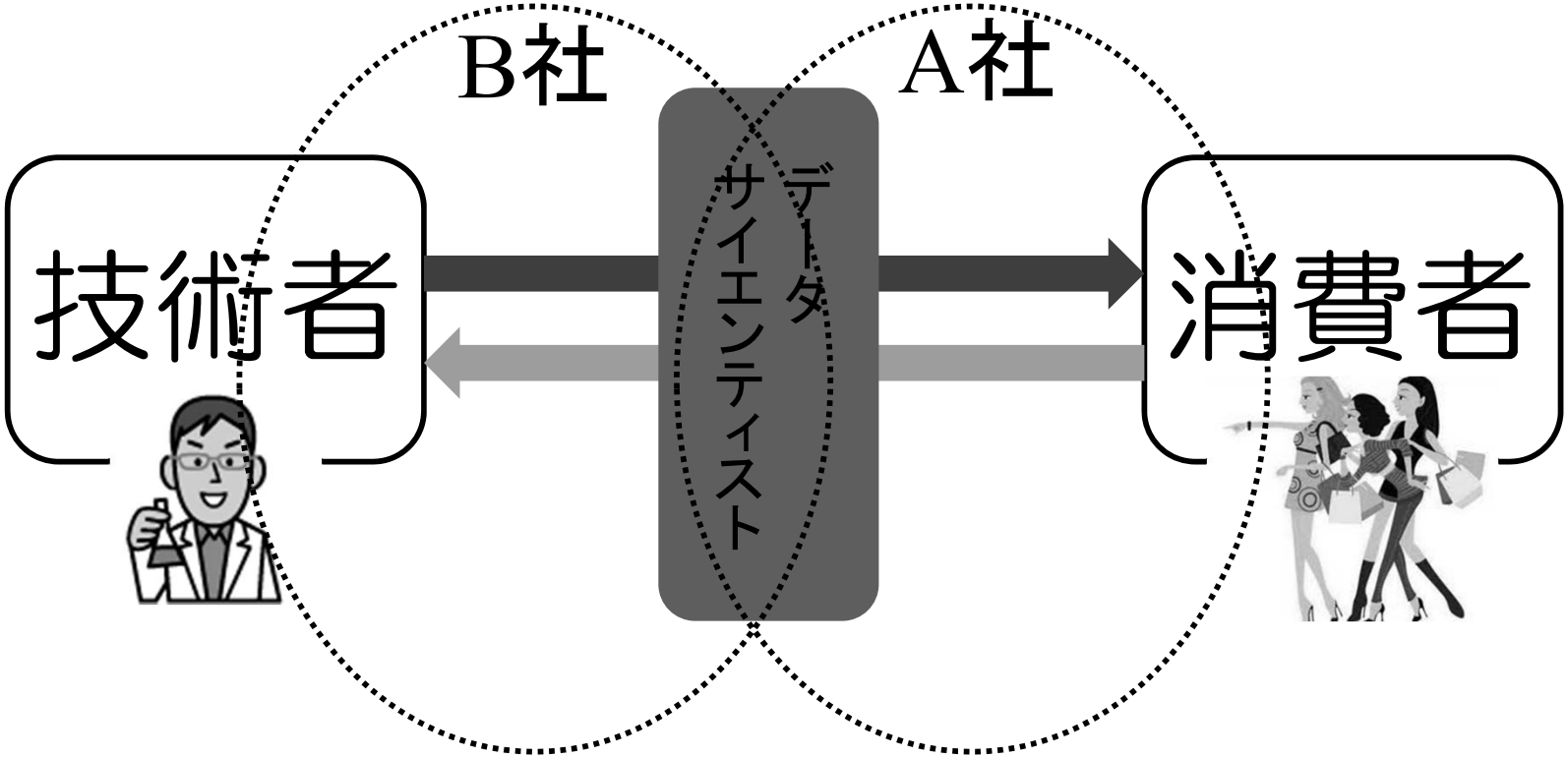


理論と仮定から結果を導く

vs.

結果から原因を探る

帰納と演繹



理論と仮定から結果を導く

vs.

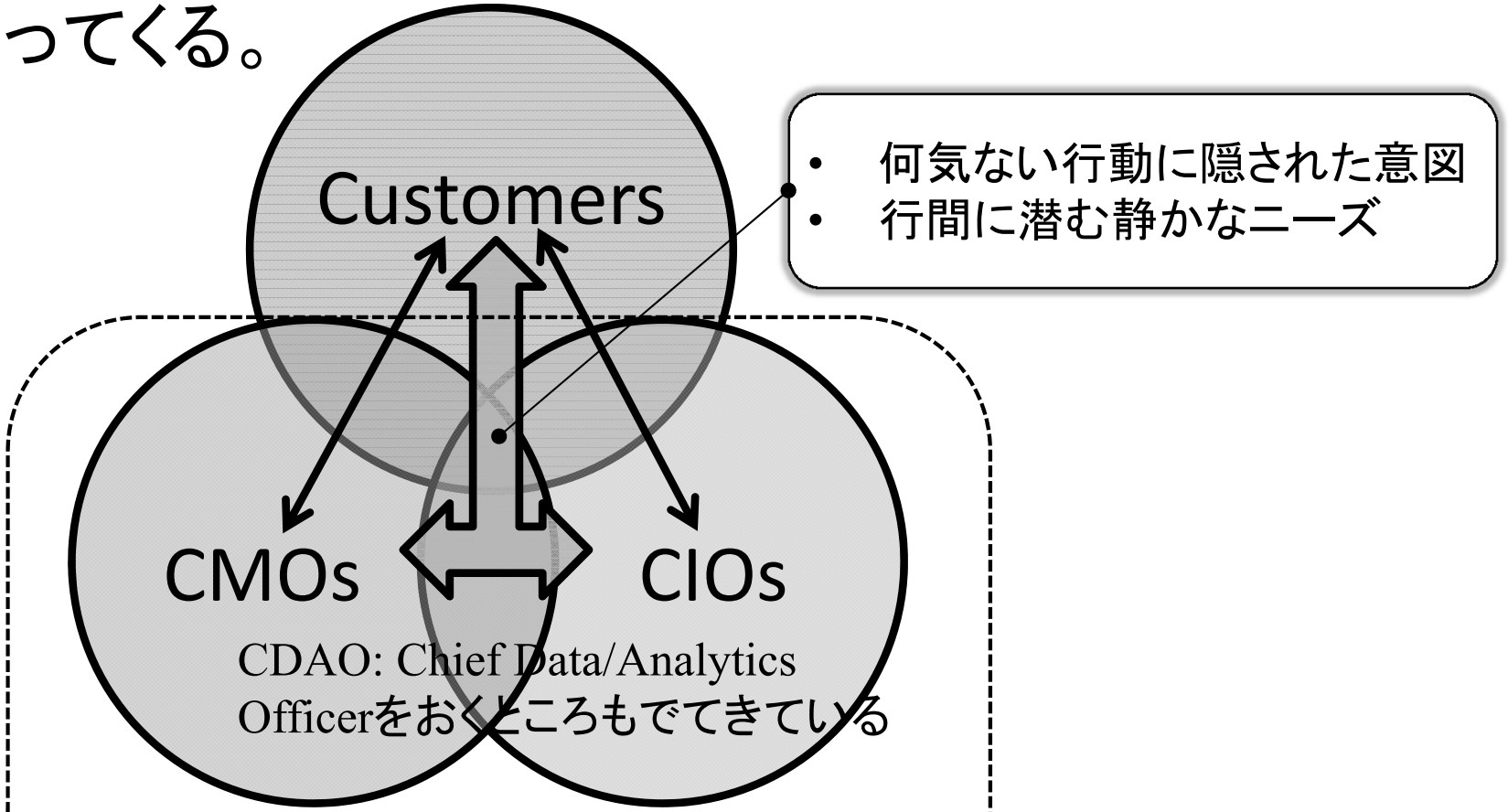
結果から原因を探る

CMOとＣＩＯの協働作業が大切

CMO: Chief Marketing Officer CIO: Chief Information

データ解析・分析の『個人商店時代』の終焉

データの価値次第でインフラやアーキテクチャの設計が大きく変わってくる。



http://blogs.hbr.org/cs/2012/11/why_cmos_and_cios_need_to_team.html

あと一言

温故知新

1993年5月：樋口助教（31才）が赤池所長（65才）に問う

(**樋口**) 最近アメリカで、ニューラルネットワークをつかったデータ解析法が非常にはやっています。ニューラルネットワーク自体はある程度脳の構造をsimulate しようというconcept はあるんですけれども、使い方として、一種の非線形回帰として使われている場合が多い様です。このニューラルネットワークモデルによるデータ解析法に対して、どの様なご意見がございますか?

(**赤池**) . . . 少し前振りのあと..

Principle があるというんなら、そういうprinciple は何を見ても解るように書かれてあるはずだけれども、どうもそういう話じゃなくて、何段階にも何段階にもlayer になってそれを繋いでうんぬんかんぬん... という話だけで。Layer になってるということは解るんだけど。感じとしては、人間の頭の中のメモリーのlayer というものは非常に大事であって、それを通じて我々の知識とかが表現されていったり、activate されたりするようになっていると思うんです。**Concept というものをどうやって作っていくかということと、記憶を使う時の構造とが対応**している。そういうところが**まだ、本質的に生きていないのではないかという気**がします。

僕自身は、統計は脳がやっている作業で、しかもそれは社会的な経験と結びついているし、情報処理と結びついているし、もちろん頭の動き。もっともっと言うと、expectation の構造と似ていると思っている。Expectation というのは、まさに昔(?)subjective Bayesianもよく言っていたんだけども、**実はその人の個性の表現**だからね。個々のケースに、**その人の持つexpectation**。そのへんを**explore していけば、無限に面白いことがある**わけ。だけど、それを一つの、大事な数学的表現として、たくさんの人の経験に基づいて、確率的な構造と対応させながら今までの経験を整理して見ていくという見方を統計は提案しているわけで。これはなかなか大きい事じゃないかなと思う。人間の中に関する事でそれだけ客観性を持った表現をして成功しているものは無いでしょう。

（『赤池弘次：統計科学を語る：1993年. 駆け出し研究者によるインタビュー記録』 by 樋口知之、川崎能典）

1993年5月：樋口助教（31才）が赤池所長（65才）に問う

さきほどの続きで、

(赤池)
前の阪大の???の所長をされていた文化勲章をもらった岡田(ヨシオ) さんが話をしていて、「いやー、ニューロ。ニューロとコンピュータの会社の人があんな話をするけど、行って勉強会で話を聞いても『別に脳神経の動きと何の関係もないんで、そんなのニューロと呼ぶのやめてくれないか』と僕は言ってんだよ」って。そういう関係の人だから、僕にペロッと冗談言った。まだそんなにそれを重視するということが必要ではないのではないか。だけども関心はもちろんあります。**大事な点があれば、我々も学ぶべきだし。この辺はそれこそ甘利先生あたりにお聞きしたら、本当はよく解る。**このあいだもヨーロッパの会議で、何かその話が出たらしくて、イギリスのstatistician が「こうやってみると統計の方が偉いんだ」という話をしたという話かなんかを、チョッとうかがったことがあるけどね。**人間の感じと結び付けていくということにいくと、これまで統計的方法が積み重ねてきている人間の経験の定式化というのは、やっぱりかなり根強く有効で、強いものがあるんじゃないか。そういうものが合流していく方向にはあるだろうとは思います。**

（『赤池弘次：統計科学を語る：1993年. 駆け出し研究者によるインタビュー記録』 by 樋口知之、川崎能典）

In memory of Prof. Akaike

2004/Nov/20