



“シン・ニホン” AI×データ時代における 日本の再生と人材育成

情報・システム研究機構シンポジウム
“分野を超えたデータサイエンスの広がり”
February 20, 2017

安宅 和人

Chief Strategy Officer
Yahoo! JAPAN



今のお仕事…



Chief Strategy Officer (CSO)

- 市場構造、変化の可視化、予見
- 全社戦略フォーカス策定
- 全社横断の経営課題解決
- ビッグデータ戦略
- 大型提携推進 など



DataScientist Society

理事

スキル委員会 委員長



理事



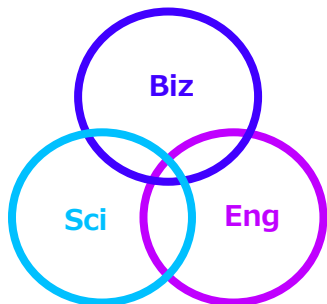
慶應義塾大学
湘南藤沢キャンパス

特任教授

「データ・ドリブン社会の創発と戦略」
with 村井純先生 + 鈴木寛先生

教育の標準化に向けた試み

スキル定義および
スキルチェックリストの整備



一般社団法人
データサイエンティスト協会

〒208-0071 東京都港区白金3-2-10 白金ビル
Tel: 03-6721-9001 Fax: 03-6721-1219

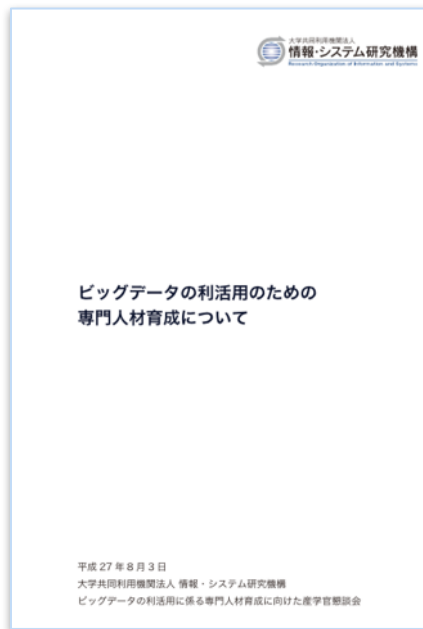
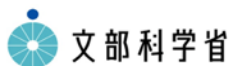
2015年11月20日

プレスリリース
報道関係各位

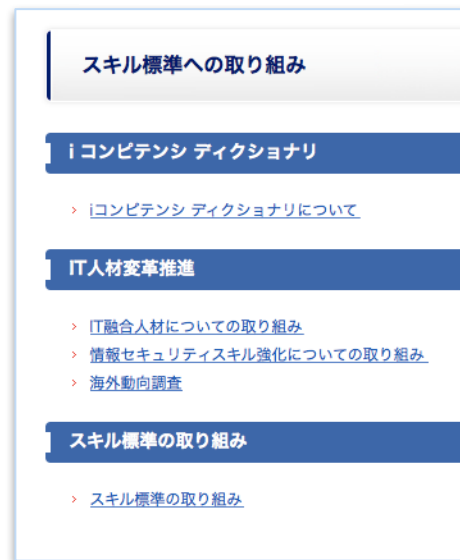
一般社団法人データサイエンティスト協会

データサイエンティストに必要なスキルをまとめた
スキルチェックリストを初公開
総合人材サービス会社での活用が決定

高等教育プログラムの検討支援



スキル標準化の検討





現在のpublic works



経済産業省

Ministry of Economy, Trade and Industry

- 経産省 産業構造審議会 新産業構造部会

- 委員

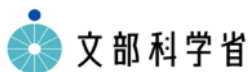


総務省

Ministry of Internal Affairs and Communications

- 人工知能技術戦略会議 産業連携会議

- 構成員



文部科学省

- 人工知能 産業化ロードマップタスクフォース

- 副主査



内閣官房

Cabinet Secretariat

- 内閣官房 第四次産業革命 人材育成推進会議

- 委員



国土交通省

- 国交省 i-Construction推進コンソーシアム

- 企画委員



Better Life with IT

- IPA スキル標準化 データサイエンス領域

- メインリード



日経調 一般社団法人
日本経済調査協議会
JAPAN ECONOMIC RESEARCH INSTITUTE

- 日経調 庄山委員会（人工知能の影響検討）

- 委員



現在のrecap

歴史的な局面

新しい
リソース

産業革命 (18~20世紀)

- 内燃機関
- 石炭と石油
- 電気工学



起きる変化

- 人間と家畜を肉体労働、手作業から開放する

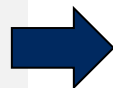


情報産業革命 (現在)

- ビッグデータ
- 高い計算能力
- 情報科学の進化



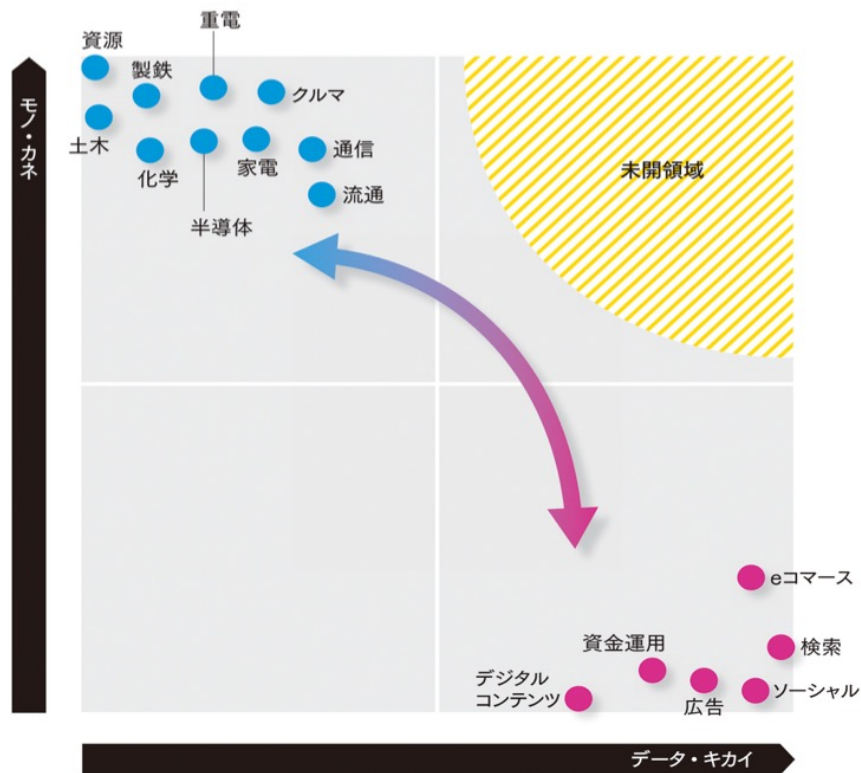
- 人間を面倒な数字入力や情報入手、情報処理作業から開放する



すべての産業がICT化する

場面	ICT化する産業
街	ファッション 小売 外食 レジャー 観光 交通サービス、       
住宅	住宅 AV機器 白物家電 住宅設備機器、     
クルマ	自動車 自動車部品 カーナビ、      
健康 医療	ヘルスケア機器 医療サービス 医薬品、      

2つの経営資源



New York City 1900





New York City 1913



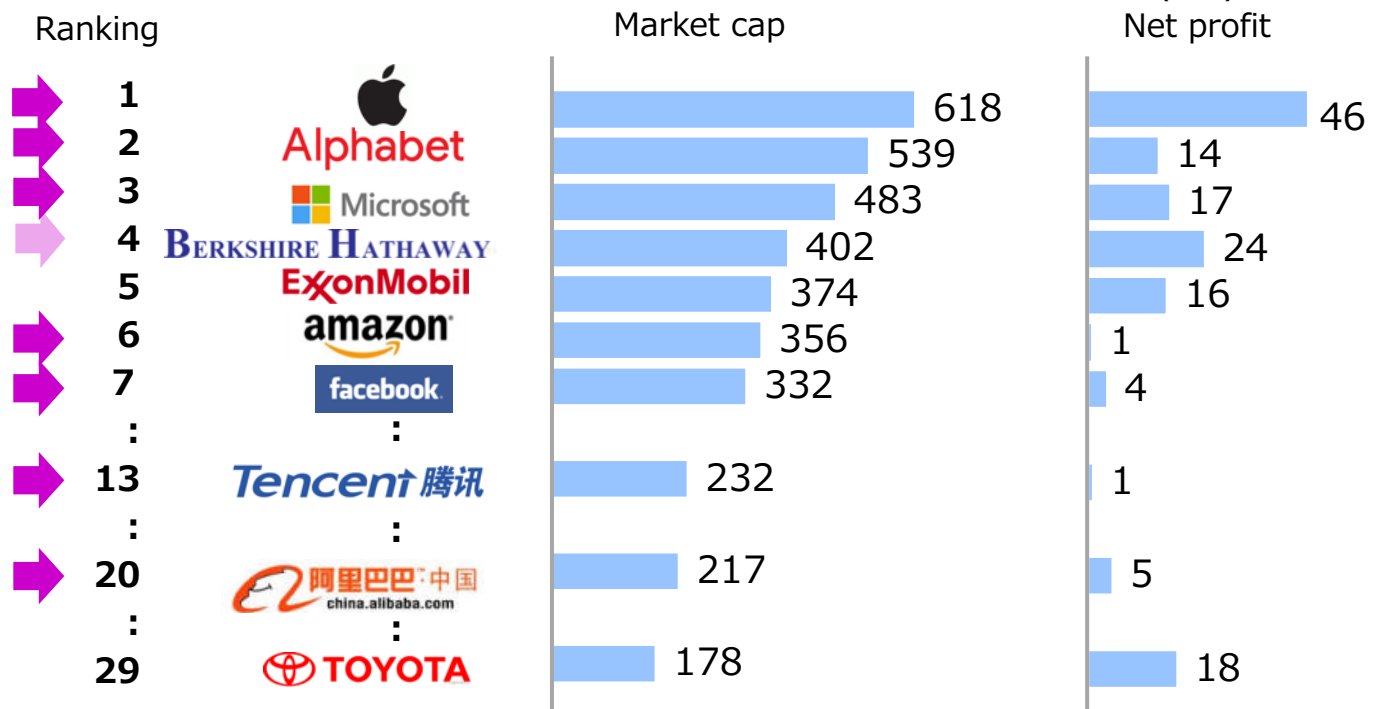
大半の人が思っているより
遙かにはやく変化は起きる



時価総額ランキング

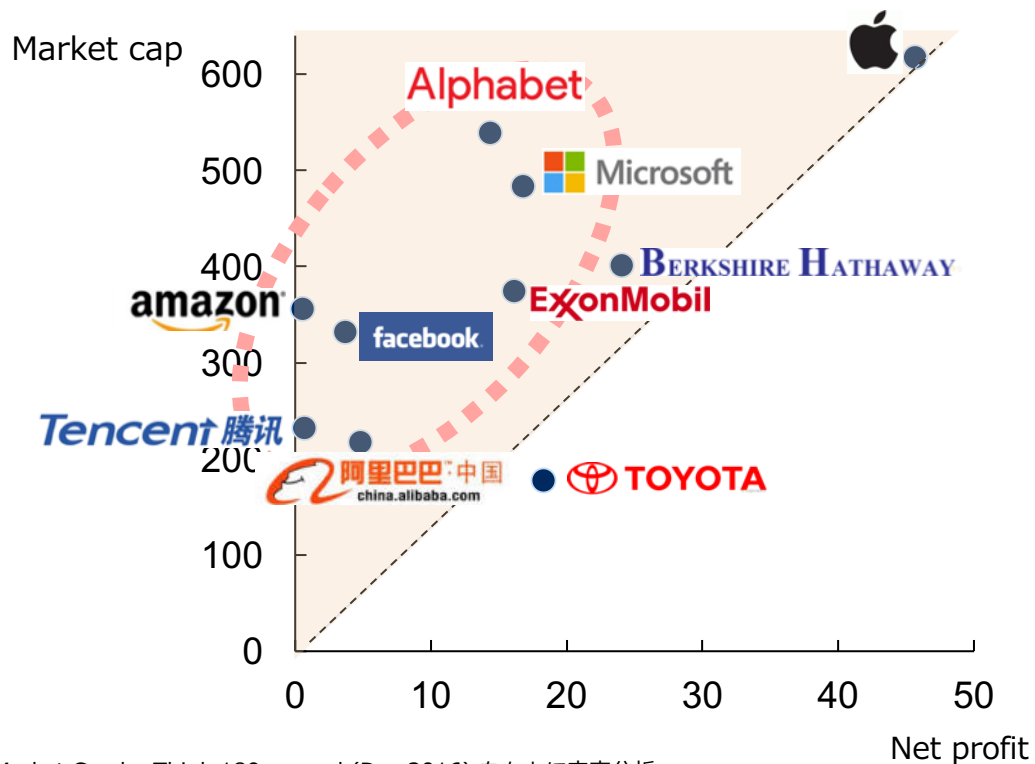
billion dollars

➡ ICT companies



時価総額と利益の関係

billion dollars



資料: World Stock Market Cap by Think 180 around (Dec 2016) をもとに安宅分析



国富を生み出す方程式が本質的に変化

Old Game

- 付加価値/GDP
- 寡占
- 既存の枠組みの中での規模と効率の追求
- 既存のルールでのサバイバル

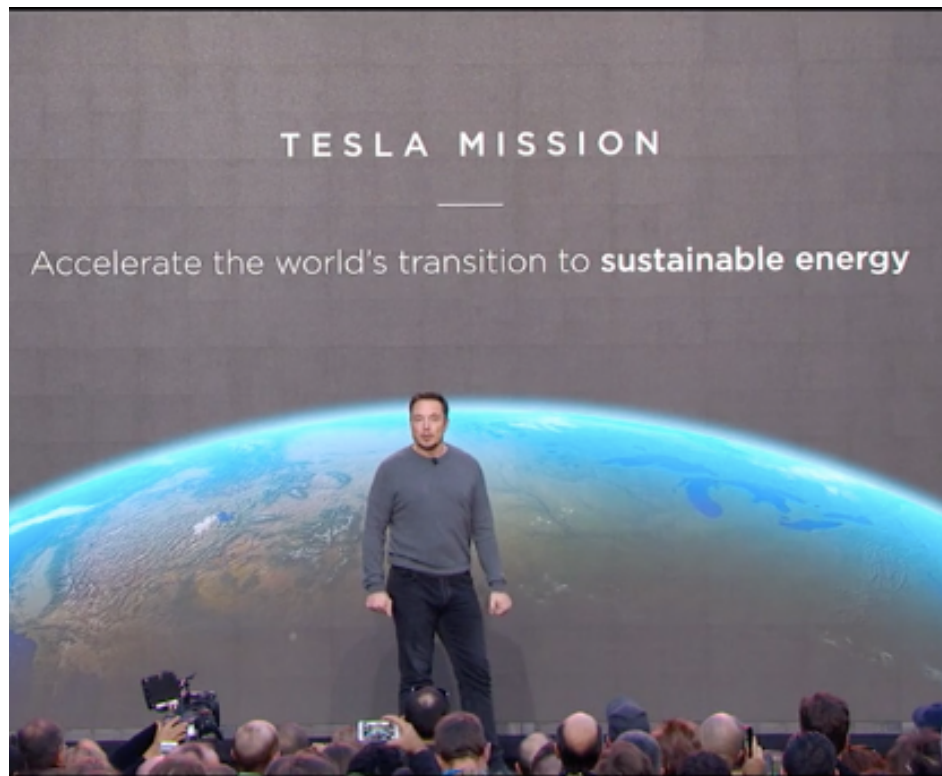


New Game

- マーケットキャップ
- 未来への期待感、寄与
- 既存の枠組みを越え、ICT、技術革新をテコに世の中をアップデート
- ジングルを切り開きサバイバル

妄想しカタチにする力が富に直結

国富の方程式の変化からの要請

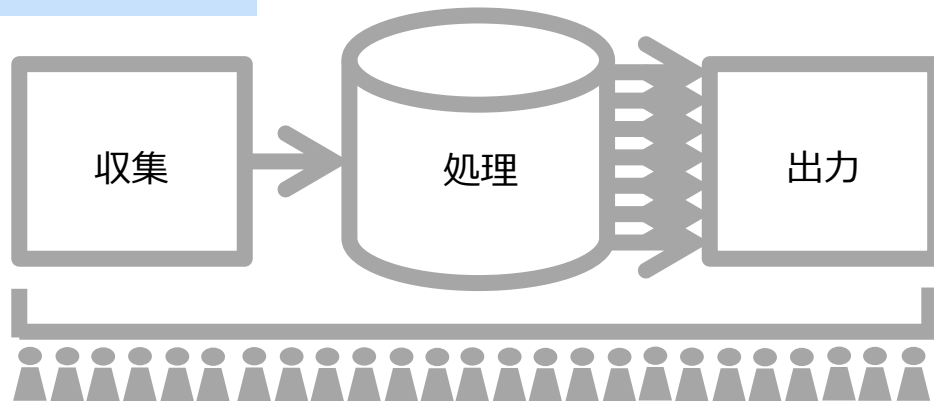


資料 : https://commons.wikimedia.org/wiki/File:Steve_Jobs_Headshot_2010-CROP.jpg
<https://www.tesla.com/powerwall?redirect=no>

AI×データ戦争における3つの成功要件

① デバイス・領域を
超えたマルチビッグ
データの利活用

② 圧倒的なデータ処理力
(データセンター・技術)

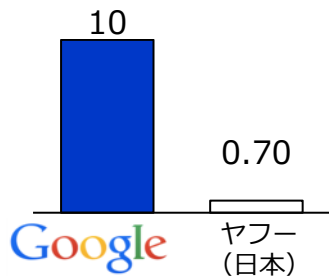


③ 質と量で世界レベルの
情報系サイエンティスト
とICTエンジニア

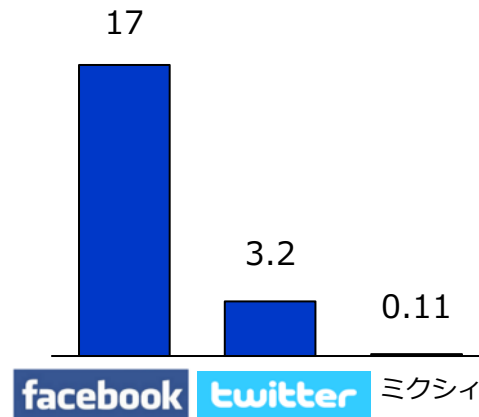
データの巨人たちとの戦い

月間利用者数 (単位: 億人)

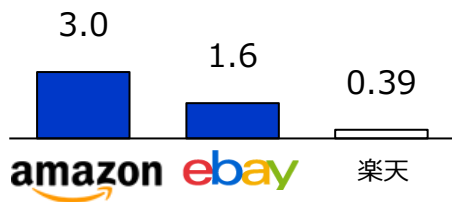
検索、ポータル



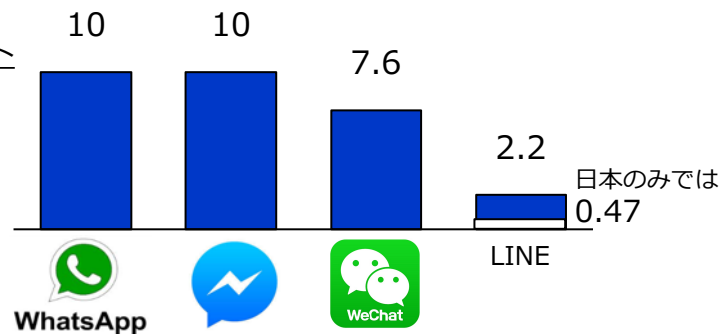
SNS系



eコマース



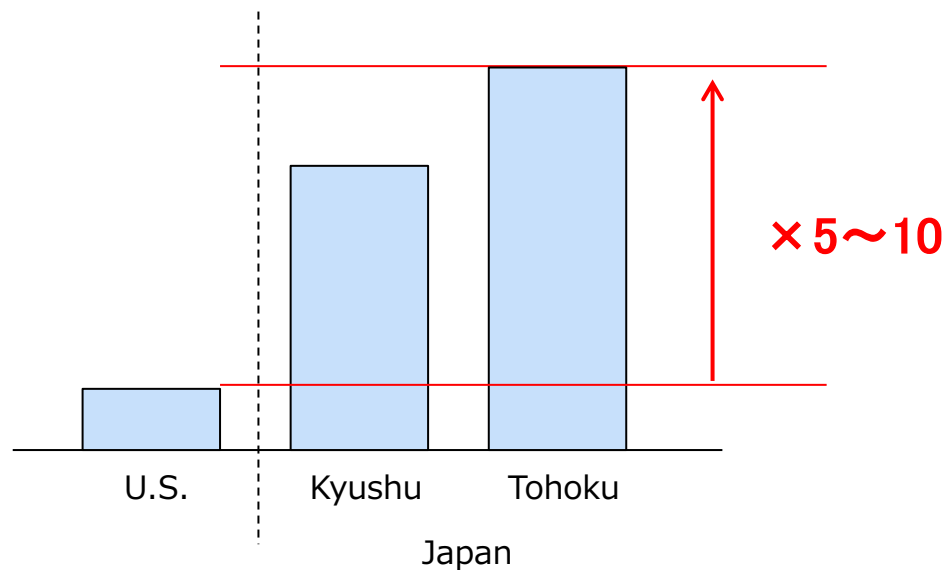
チャット





データ処理コスト比較

cost of industrial electricity (yen/kwh)





ビッグデータ技術のほとんどは海外





深層学習…すべての主要拠点は国外

NYU
(Yann LeCun)



U Tronto
(Geoffrey Hinton)



U Montreal
(Yoshua Bengio)



Stanford U
(Andrew Ng)



英国 DeepMind
(Demis Hassabis)

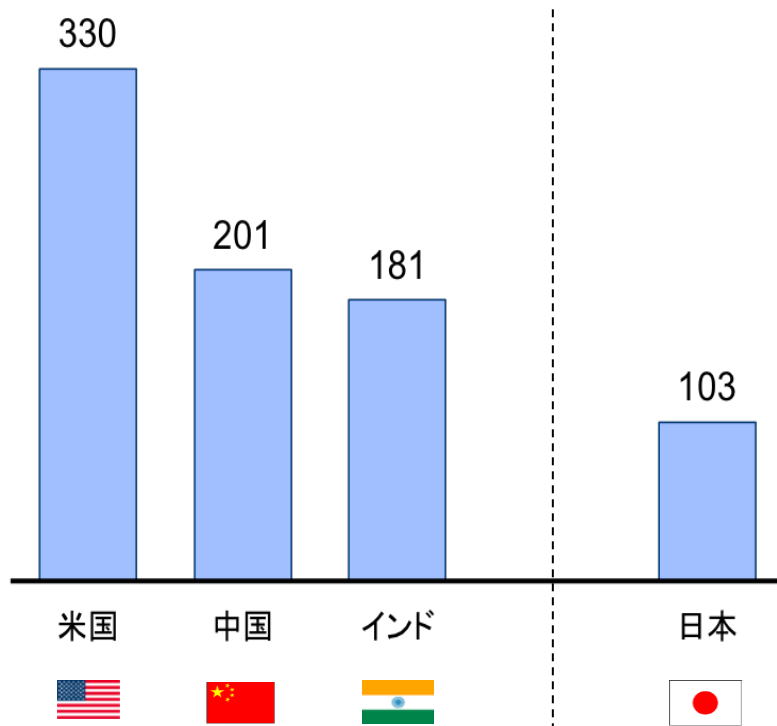


IDSIA
(Jürgen Schmidhuber)



人材数自体に大きな課題

ICTエンジニアの数の国別比較：万人





エンジニアリング層の現状

- 大半がSIerにおけるcoderというべき人材
- 研究と開発のギャップを乗り越えられる人*が少ない
- 高速データ収集、分散環境、ロギング周りの仕組みを作れ、回せる人が極めて限定的

*研究・実験環境を堅牢で大規模かつリアルタイムの本番環境につなげられる人材

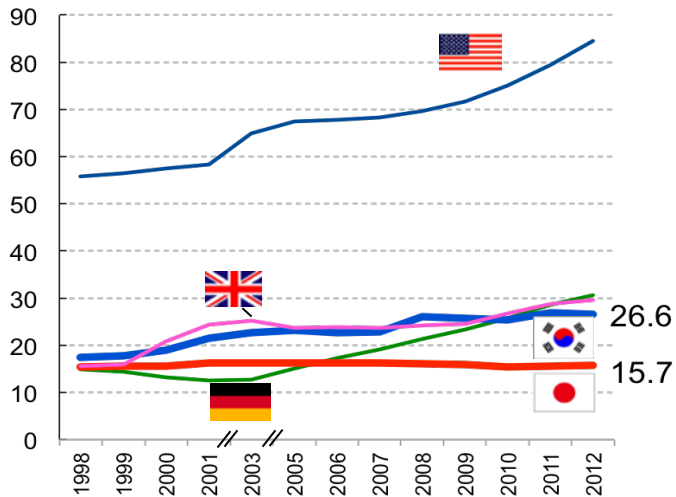
資料: 安宅和人「データ時代に向けたビジネス課題とアカデミアに向けた期待」応用統計学セミナー2015.5.23 (<http://www.applstat.gr.jp/seminar/ataka.pdf>)



理工系の学生の数自体が足りない

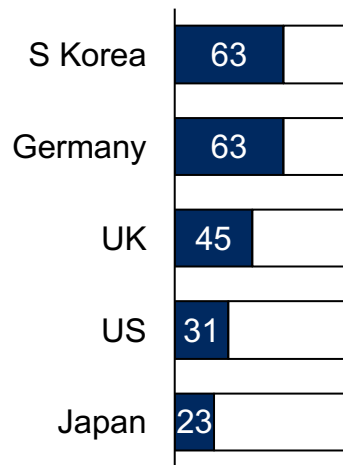
Number of science and technology graduates

(10k/year)



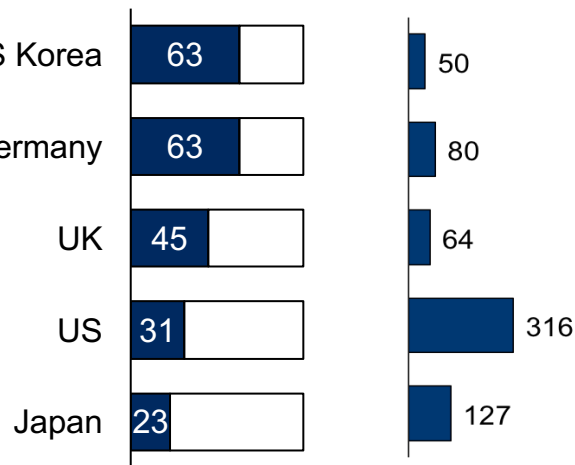
Ratio of science and tech major student

(% 2012)



Population

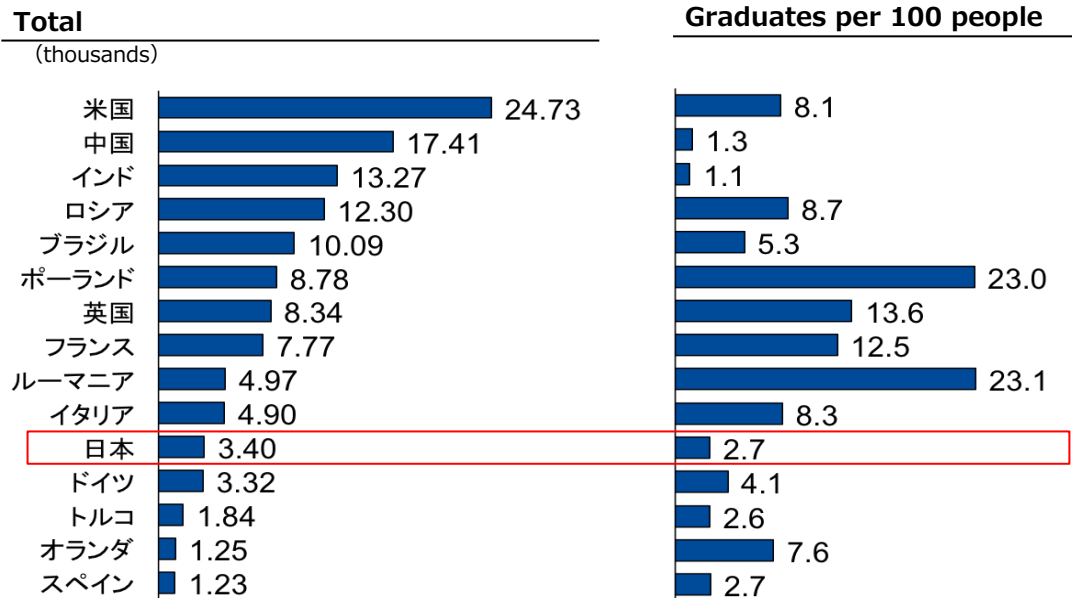
(millions)



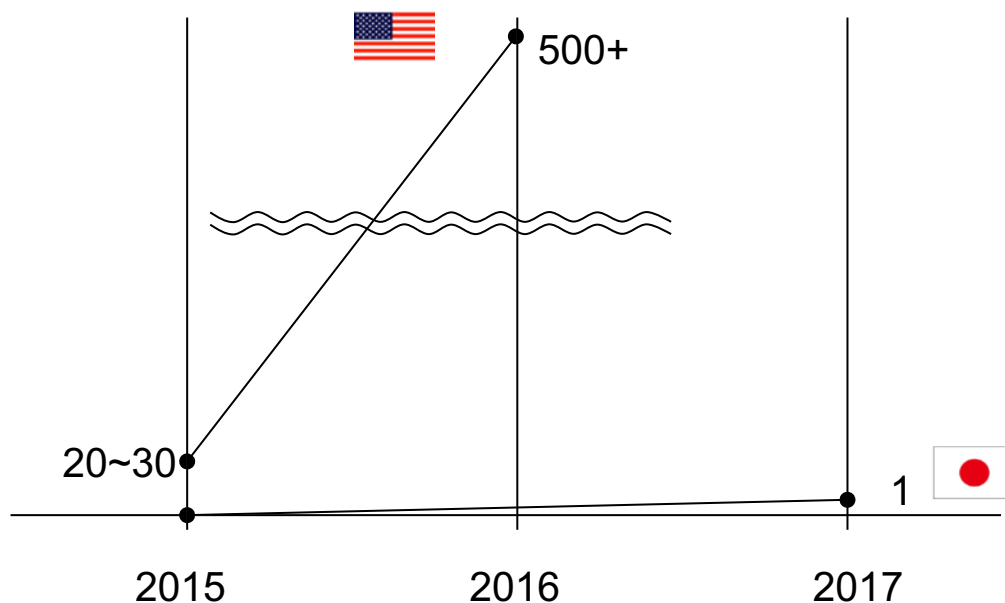
※理工系：工学、科学、数学、物理など
(医学、薬学は含まず)



深い分析訓練を受けた大卒の数も少ない



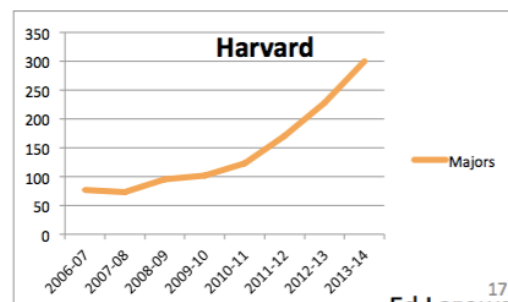
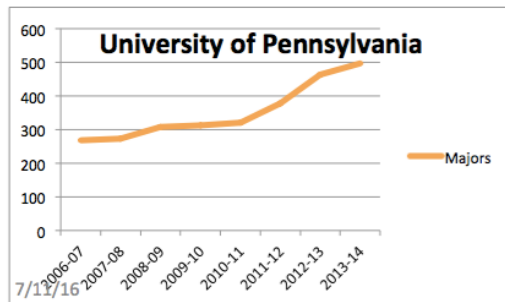
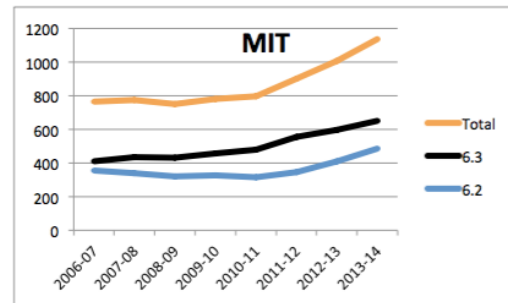
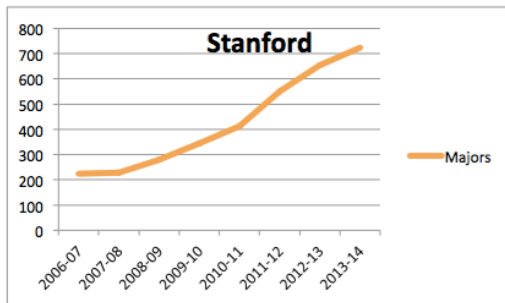
Number of Data Science degree programs





海外トップスクールでは計算機科学はデフォルト化

of undergraduate computer science major in US top schools

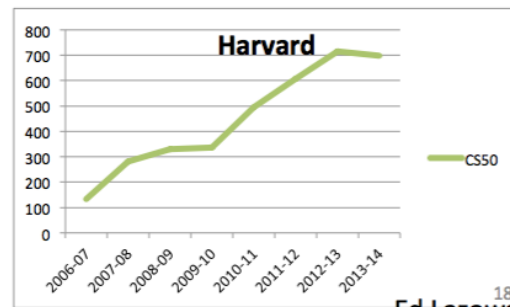
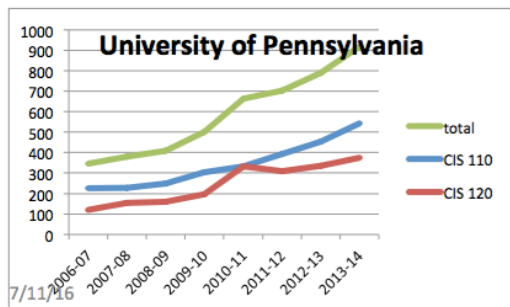
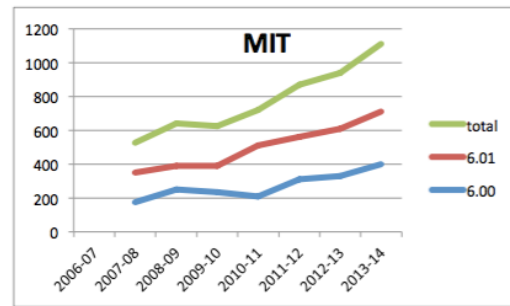
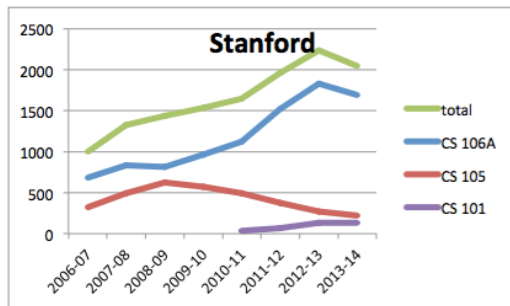


Ed Lazowska



初級的コース履修者は爆発的增加

of introductory computer science enrollment in US top schools



Ed Lazowska

我が国の新卒層の課題

基本的な問題解決
能力の欠落

- 問題を定義できない
- 結論を出すことができない

数字のハンドリン
グの基本が欠落

- 指数と実数の使い分けができない
- 指数を指数で割ったりする

分析の基本ができ
ていない

- 数字を並べることと分析の違いがわかっていない
- 軸を立てるということの意味がわかっていない

基礎的な統計的素
養がない

- 平均を鵜呑みにする
- サンプルング、統計的な有意性の概念の欠落

情報処理、プログラミングについての基本的な理解が
ない



高等教育を
受けたはずの人が
基本的な
サバイバルスキル
を身につけて
いない



日本の若者たちは
持つべき武器を持たずに
戦場に出ていっている

サイエンス層・専門家*層の現状

- そもそもいない
- どこにいるのか分からない
- いても実社会での利用に関心のある人が少ない



- 供給強化だけでは不十分
- 内向きのオタクではなく世界を変えようとする geek/hackerが必要！

*機械学習、言語処理、画像処理、音声処理、データ可視化など

資料: 安宅和人「データ時代に向けたビジネス課題とアカデミアに向けた期待」応用統計学セミナー2015.5.23 (<http://www.applstat.gr.jp/seminar/ataka.pdf>)



シリコンバレーの創業者たち

YAHOO!



Jerry
Yang

Stanford
BS/MS
Electrical
engineering

Google



Larry
Page Sergey
 Brin

Stanford
PhD program
Computer
science



Andy
Rubin

Utica College
BS
Computer
science

facebook



Mark
Zuckerberg

Harvard
BS
Psychology &
Computer
science

TESLA



Elon
Musk

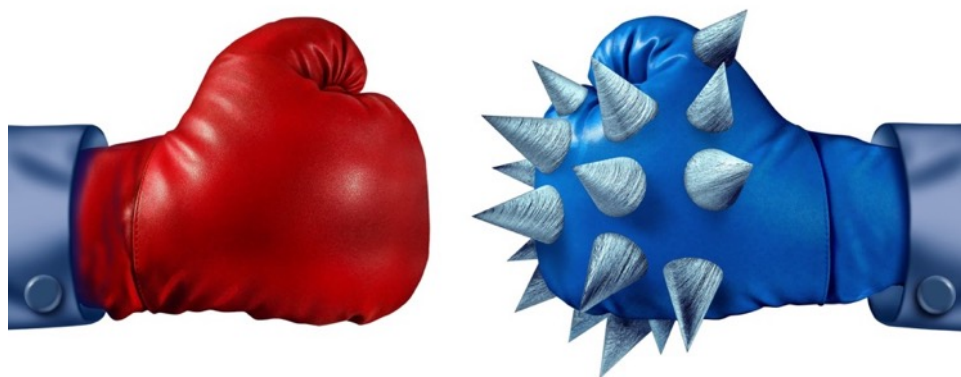
Stanford
PhD program
Applied
physics

ミドル層・マネジメント層の現状

- そもそものチャンスと危機、現代の挑戦の幅と深さを理解していない
- ビジネス課題とサイエンス、エンジニアリングをつなぐアーキテクト的なヒトがいない
- 生き延びるためにはスキルをrenewしなければいけないが、身につける方法がわからない上、学ぶ場がない



このままでは
「じゃまオジ」
だらけの社会に



Not even a competition
勝負になっていない

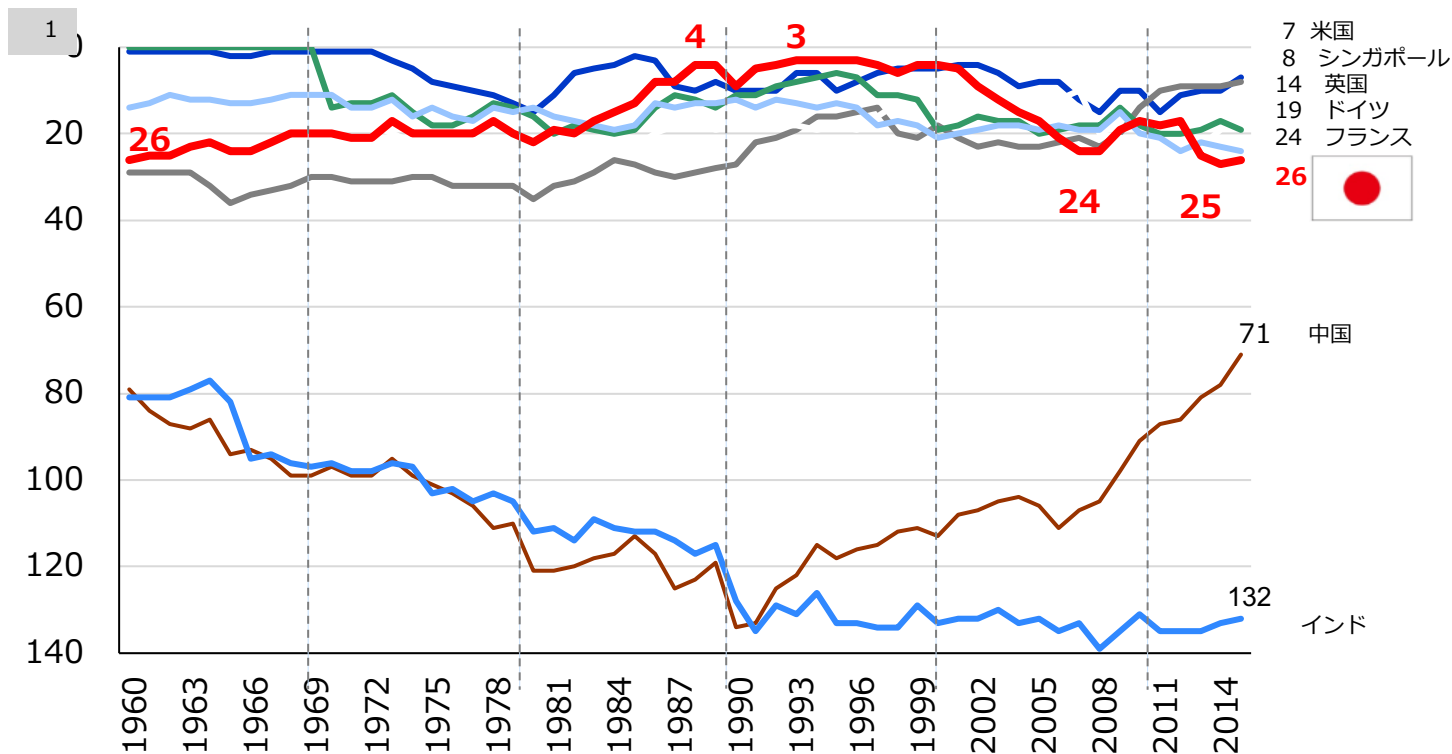
Like 164 years ago
164年前と同様





1960年代以来の水準

GDP per capita Ranking
(unit=current US\$)



GDP per capita is gross domestic product divided by midyear population.

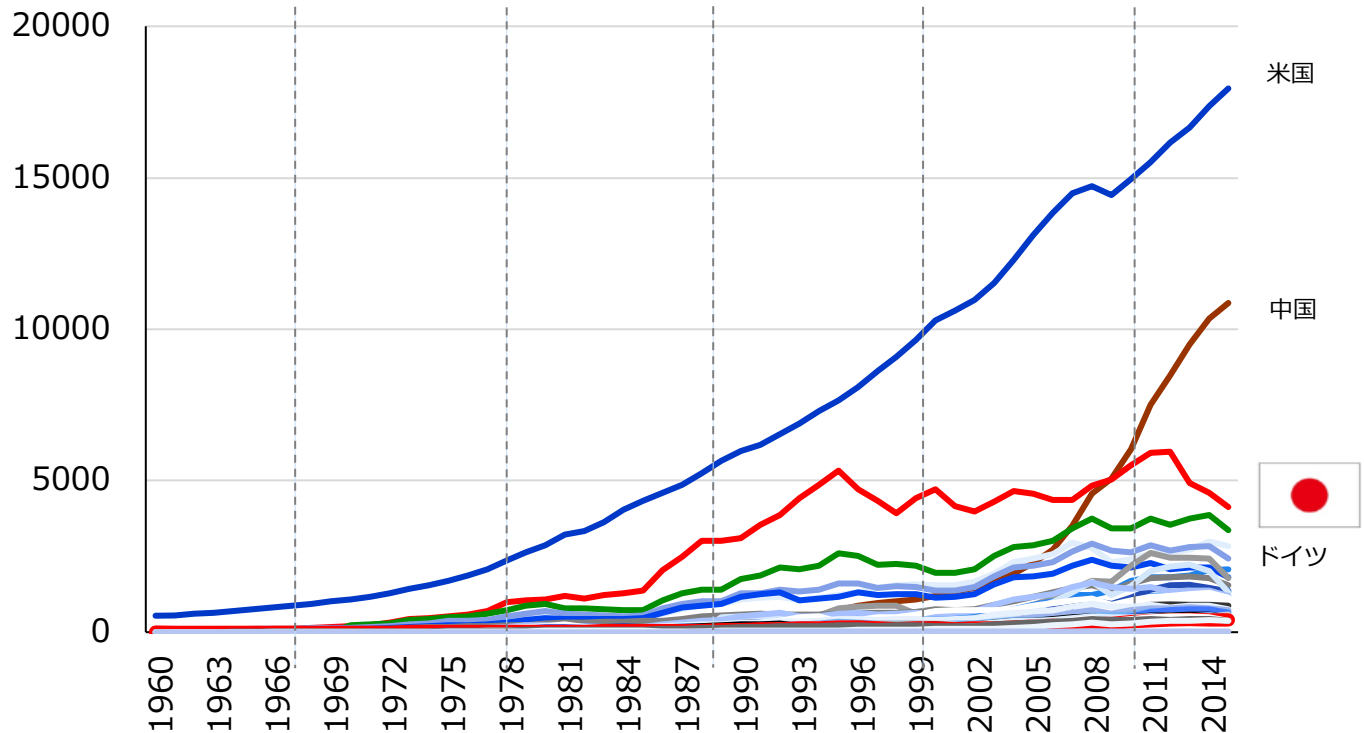
Source: World Bank national accounts data, and OECD National Accounts data files. Excluded countries which data is not available in 2015.



5年でドイツに並ばれかねない状況

Trend of GDP

(Top 30 countries in 2015, unit=current billion US\$)



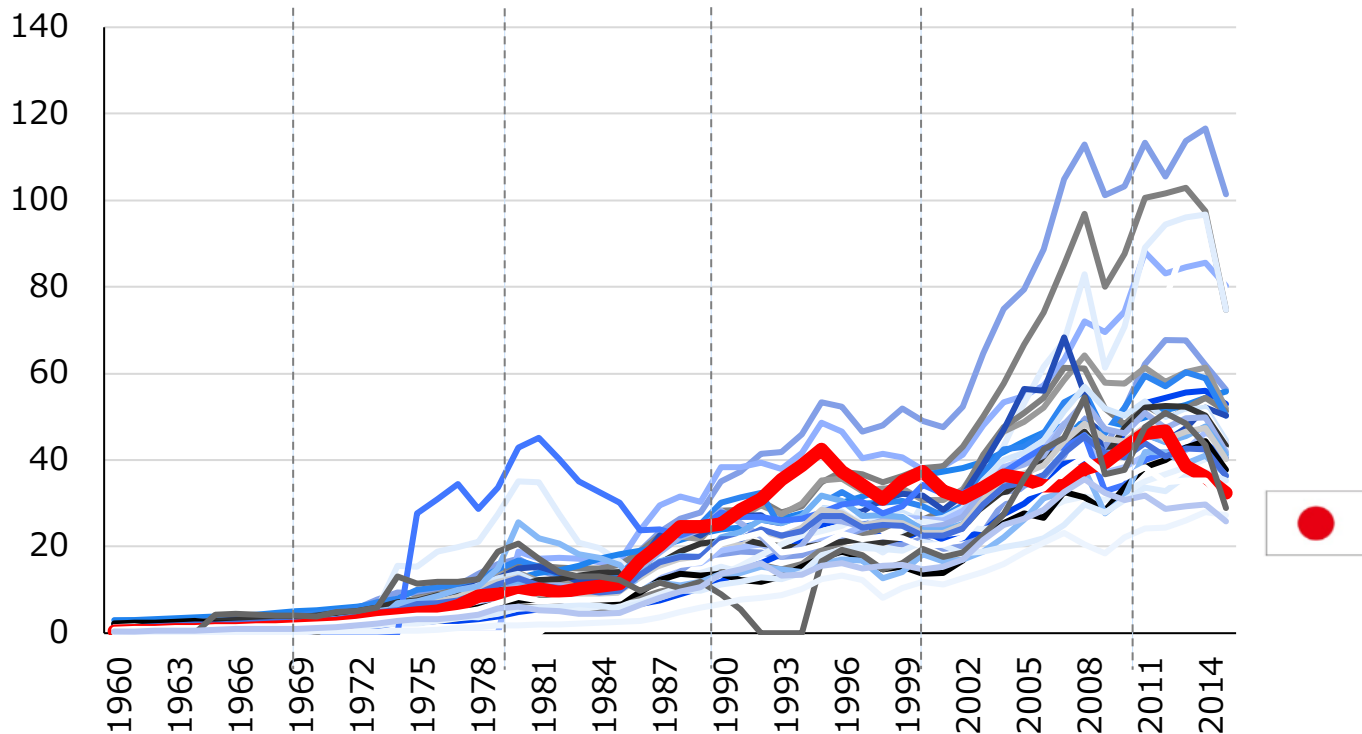
GDP is at purchaser's prices

Source: World Bank national accounts data, and OECD National Accounts data files. Excluded countries which data is not available in 2015.

半ば一人負け

Trend of GDP per capita

(Top 30 countries in 2015, unit=current 1,000 US\$)



GDP per capita is gross domestic product divided by midyear population.

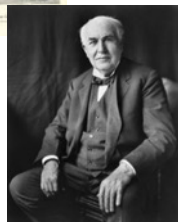
Source: World Bank national accounts data, and OECD National Accounts data files. Excluded countries which data is not available in 2015.

産業革命の三段階（大局観）

recap

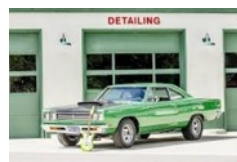
1750~

新エネルギーと技術



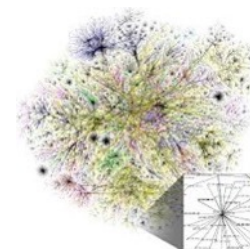
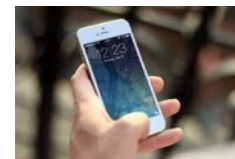
1900~

高度な応用



1960~

エコシステム構築



日本は第二の波から参加

1750~

新エネルギーと技術



1900~

高度な応用

Panasonic

SONY

TOYOTA

HONDA

Nikon

SEIKO



1960~

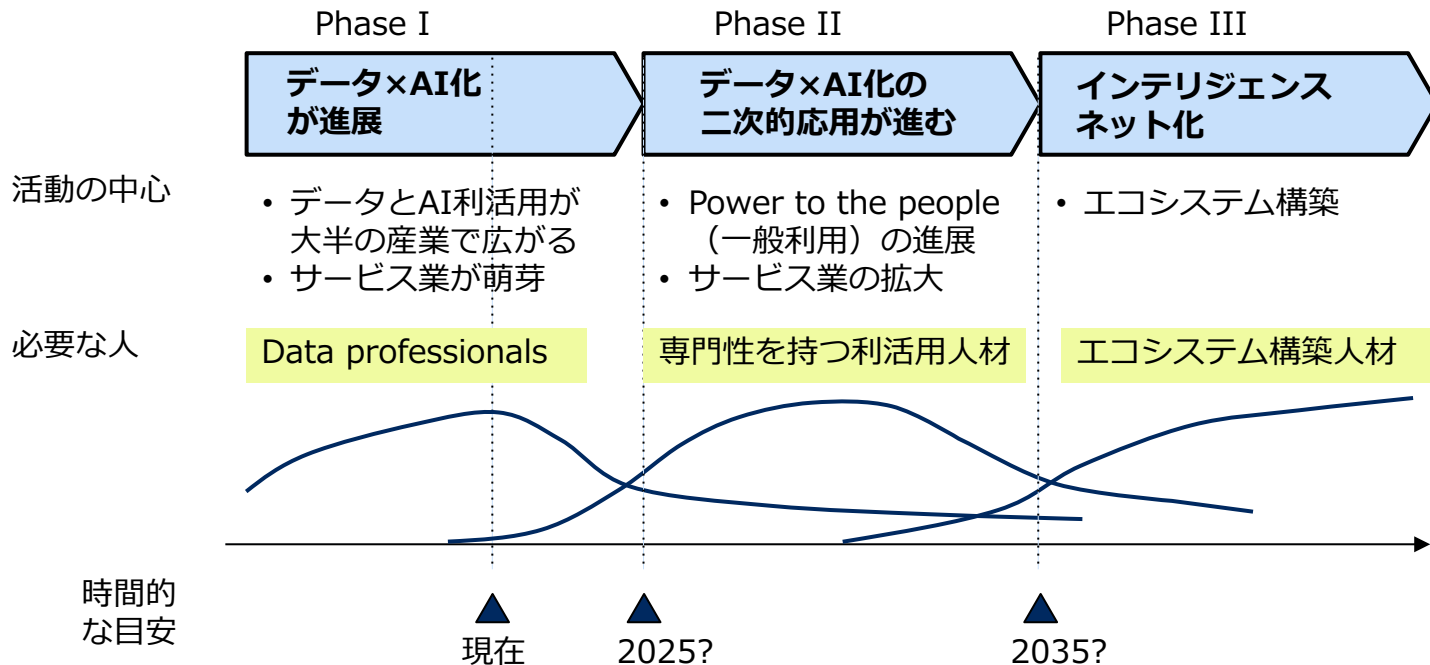
エコシステム構築



第二、第三の波に備える必要がある

データ×AI化における産業化の大局観

recap





この国はスクラップ&ビルドで
のし上がってきた。
今度も立ち上げられる。

赤坂秀樹
内閣官房長官代理
(『シン・ゴジラ』より)



どんな人が必要になるのか？



これから起きる本当の競争

よくある誤解

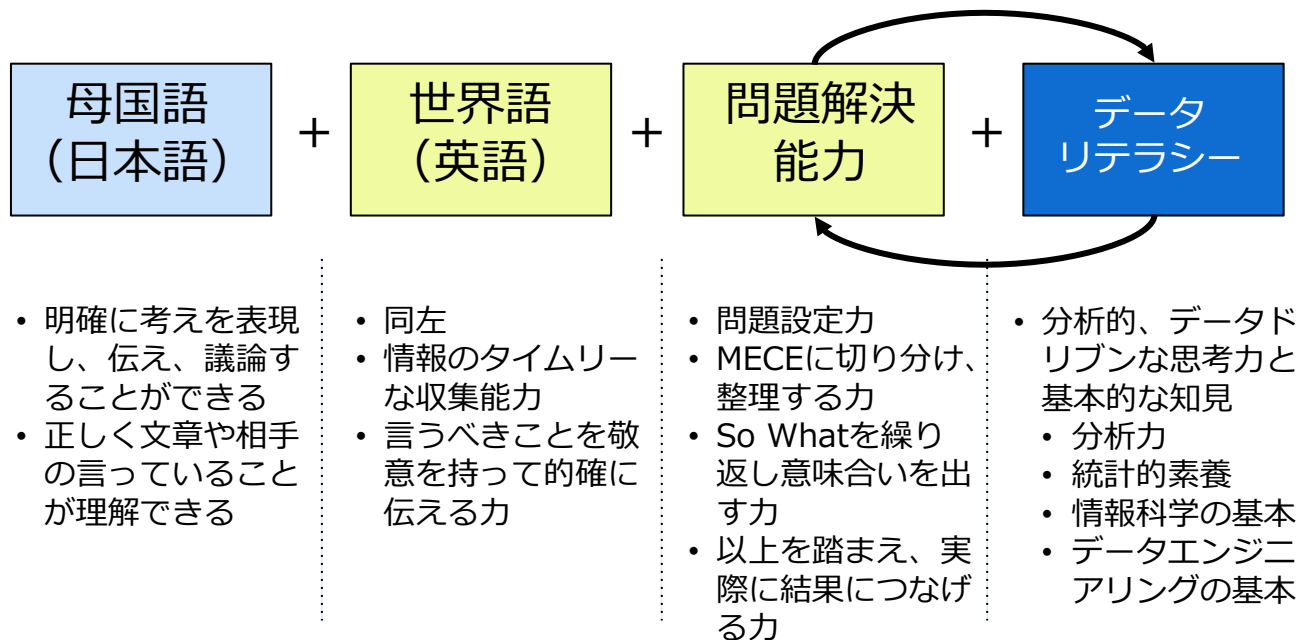
- AI vs 人間

本当の姿

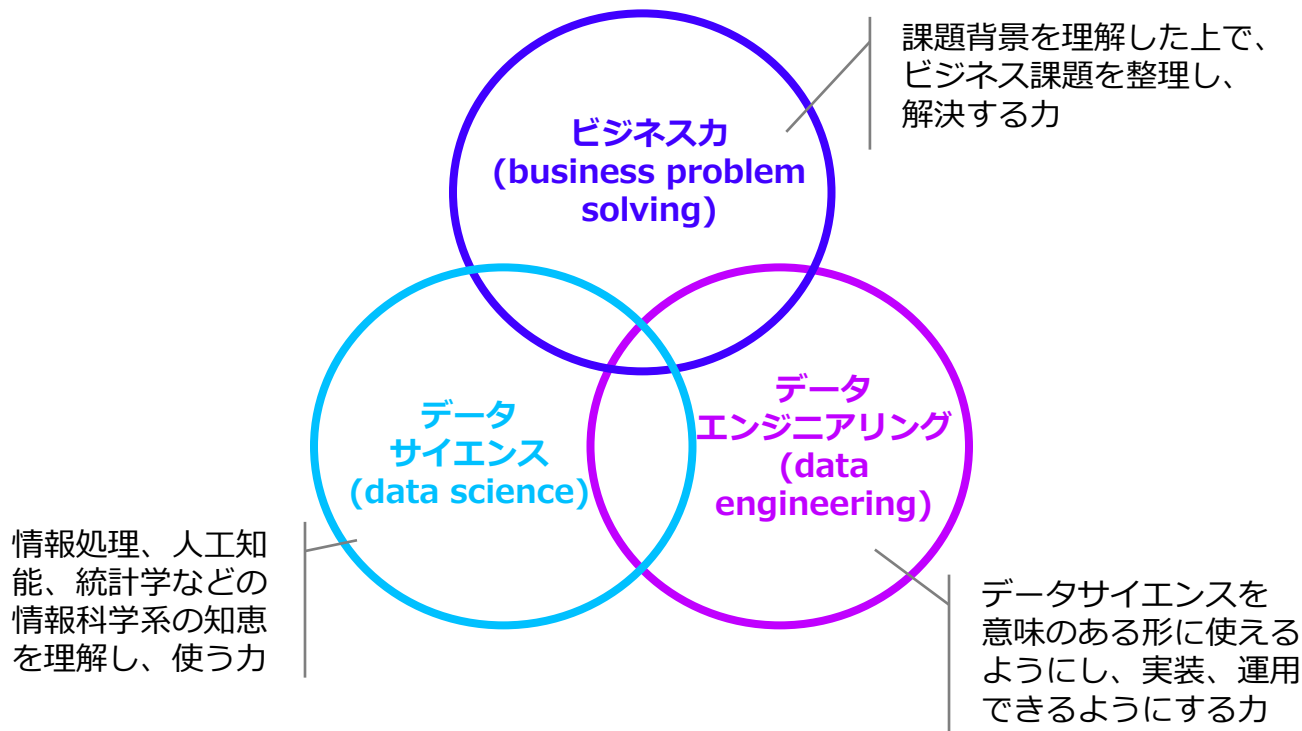
- 自分とその周りの経験たゞ
けから学び、AIやデータ
の力を使わない人
- VS
- 手に入る限りのあらゆる
データからコンピュー
ティングパワーを利用して
学び、その力を活用する人

社会を生き抜くための基礎教養が変化

現代のリベラルアーツ



データの力を解き放つための 3つのスキルセット



これまでとは似て非なるdata professional人材が必要

Not this

- 基礎研究にしか関心がない人
- 統計だけの専門家
- 単なるデータの専門家
- ただ仕様書に基づきcodingをするSE、プログラマー

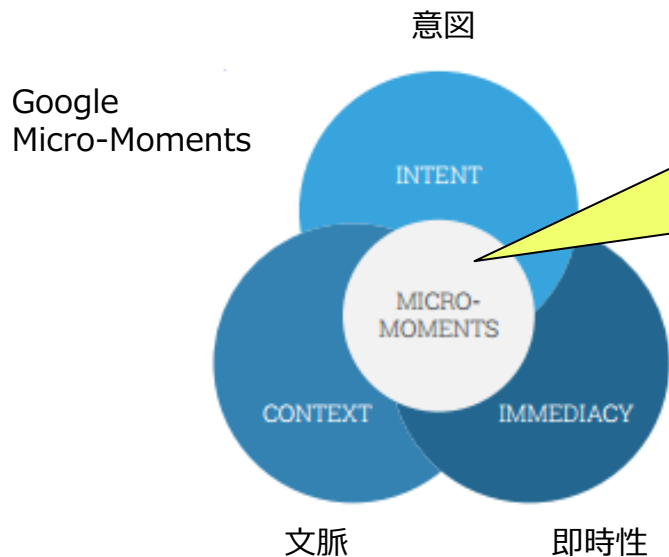


But this

- 時代の変化から生まれるリアルな課題解決にエキサイトする人
- 統計的素養を持った上で情報科学的な知恵と技を上での課題解決に使う人
- 課題を俯瞰し柔軟にビッグデータ処理を実験環境から本番環境まで実現できる人

境界・応用領域にこそ人材が必要(1)

ニーズが発生した瞬間を検出し、打ち込む機能



- 広告であっても、その人が必要な時に提示することができれば、その「広告」は非常に有用な「情報」になる

境界・応用領域にこそ人材が必要(2)

高度な熟練を要する皮膚がん診断の自動化

Dermatologist-level classification of skin cancer with deep neural networks

Andre Esteva^{1*}, Brett Kuprel^{1*}, Roberto A. Novoa^{2,3}, Justin Ko², Susan M. Swetter^{2,4}, Helen M. Blau⁵ & Sebastian Thrun⁶

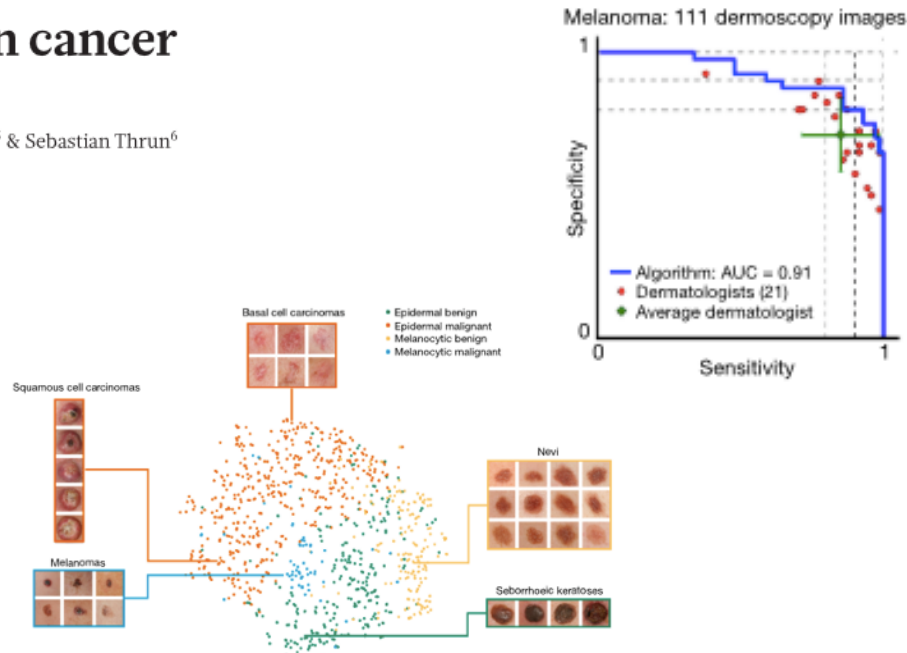
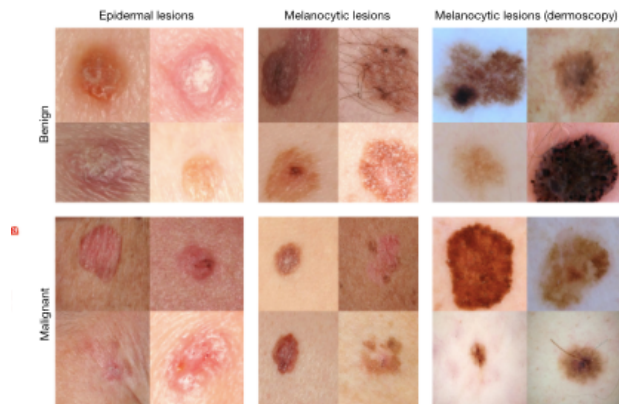
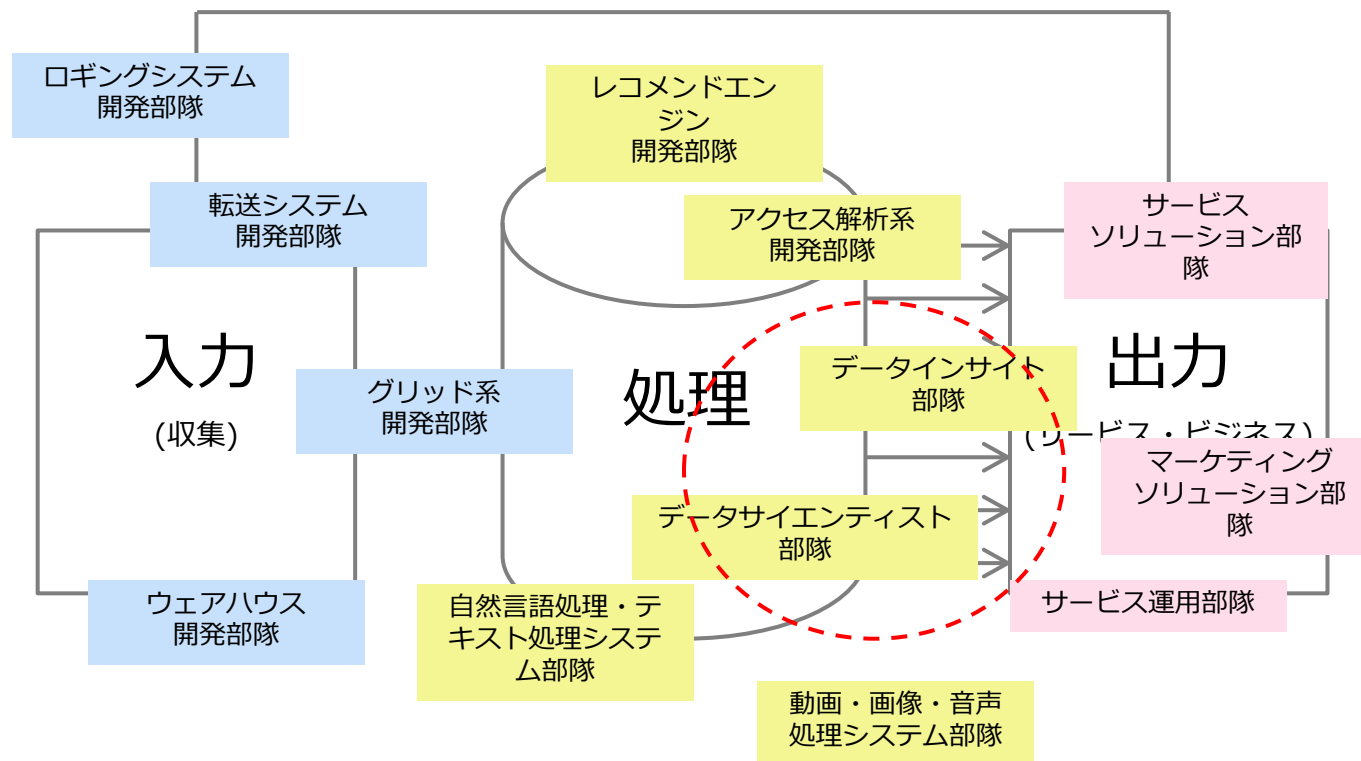


Figure 4 | t-SNE visualization of the last hidden layer representations in the CNN for four disease classes. Here we show the CNN's internal representation of four important disease classes by applying t-SNE, a method for visualizing high-dimensional data, to the last hidden layer representation in the CNN of the biopsy-proven photographic test sets

(932 images). Coloured point clouds represent the different disease categories, showing how the algorithm clusters the diseases. Insets show images corresponding to various points. Images reprinted with permission from the Edinburgh Dermofit Library (<https://licensing.eri.ed.ac.uk/i/software/dermofit-image-library.html>).

Data Scientist以外のビッグデータ人材も必要

ビッグデータ利活用に必要な専門人材の広がり



その上での見る力、決める力、伝える力が大切になる

意味合いとこれからの共存

起きる変化と意味合い

- 人間が本来拠って立つ役割が赤裸々に
- 人間は人間らしい価値を提供する事に集中することに



これからの共存

ヒト

- 総合的に見立てる
- 方向を定める
- 問いを立てる
- 組織を率いる
- ヒトを奮い立たせる



データ
×AI

- 識別
- 予測
- 実行



ほぼ全ての革新は若者によって行われてきた

初特許（創業）1868

相対性理論 1905

吉田松陰
1859

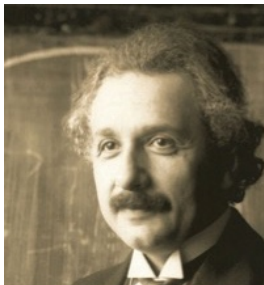


明治維新の
思想的指導者
29歳没



Thomas Edison 21歳

電話の発明 1875



Albert Einstein
26歳



Graham Bell 28歳

東通工創業 1946

SONY

井深大 38歳
盛田昭夫 25歳

Google創業 1998



Larry Page 25歳
Sergei Brin 25歳

Apple創業 1976



Steve Jobs 21歳
Steve Wozniak 26歳

松下電器創業 1917

**National
Panasonic**

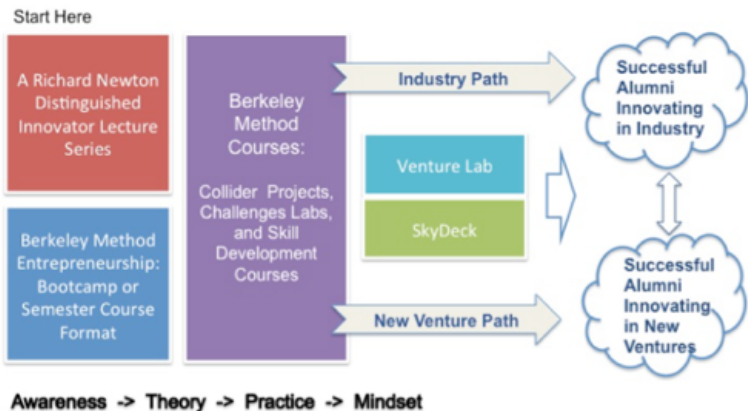
松下幸之助 24歳
井植歳男 16歳



Entrepreneurship教育も高校、学部からデフォルトで行うべき



Berkeley Method Curriculum



Stanford University



ARE YOU EXTREME?

What do you get when you cross a brain

A TRAIN STORM!

NOT YOUR AVERAGE CLASS

students + ideas + empathy + perspiration = innovation + impact

DESIGN FOR EXTREME AFFORDABILITY

Exponential thinkingが基本

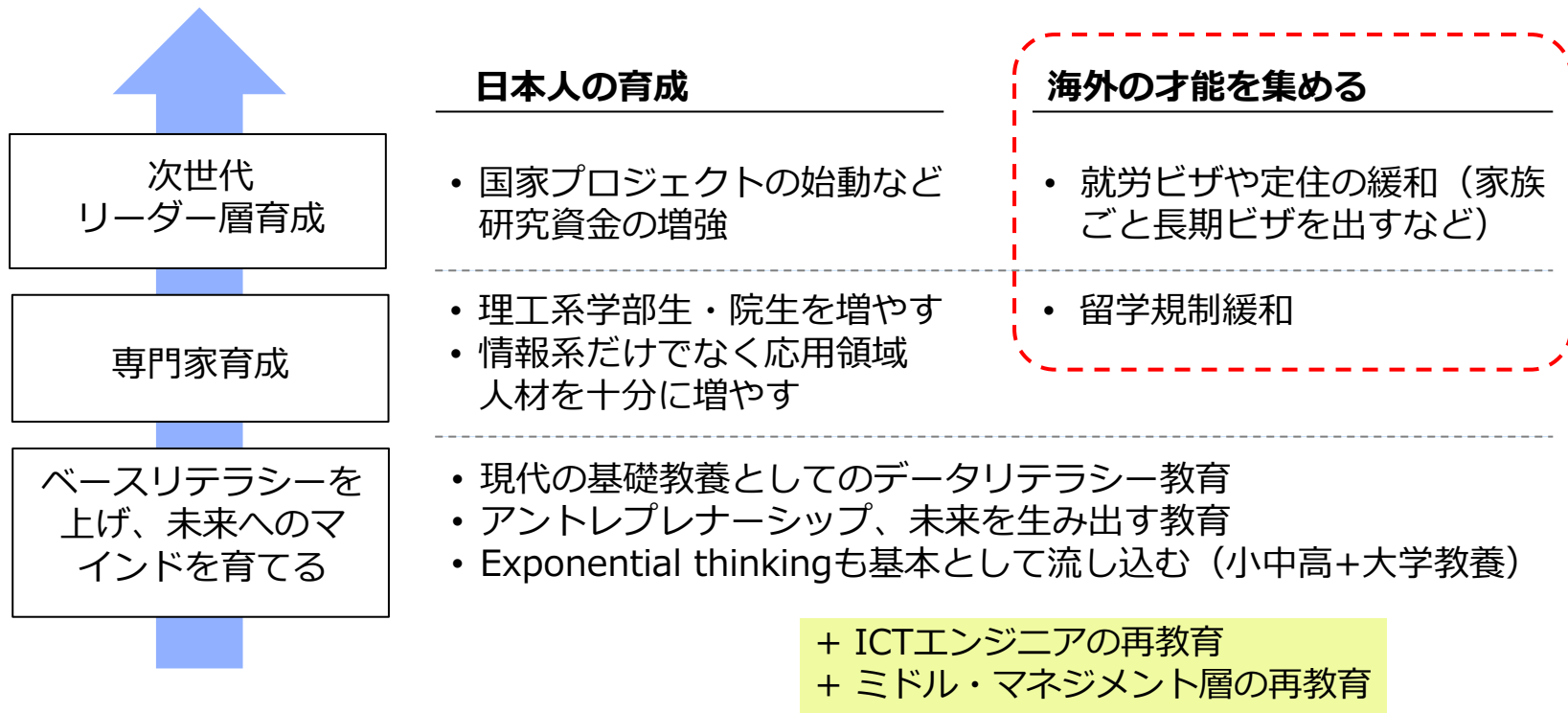
The revolution in AI
has been profound,
it definitely surprised
me, even though I was
sitting right there.

Sergey Brin
Google co-founder



3層+2で育て、加えて世界の才能を取り込むべき

AI×データ時代に向けた人材の増強イメージ



千載一遇のタイミングを活かすべき



Hiro Ono / 小野雅裕

@masahiro_ono

 Follow

#NASAの日常

怒怒怒怒怒怒怒怒

インドから超優秀な学生をインターンで採用しようとしたら、突然外国人の有給インターンは今後雇えないという通達が来た。理由は言われず。「上層部の決定」らしい。完全に推測だけど、あいつのせいだろうな。本当に悲しい。こうしてアメリカは才能を失っていく。

 Translate from Japanese

RETWEETS

2,014

LIKES

1,296



5:09 AM - 2 Feb 2017



Hiro Ono / 小野雅裕

@masahiro_ono FOLLOWS YOU

NASA JPLで火星ローバーの自動運転アルゴリズムを作ってます。宇宙探査の連載『一千億分の

八』：koyamachuya.com/column_category... ミーちゃんのパパ。阪神ファン。

 Pasadena, CA

教える人が足りなければ連れてくるのが基本

仏教の導入時
(8世紀、奈良時代)



仏教の三宝（仏法僧）をもち
世界に認められる国家になる

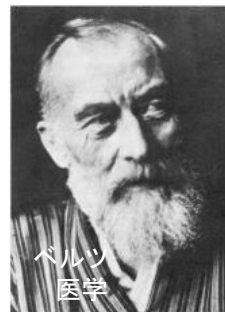
明治の開国
(19世紀)



コンドル建築学



メツケル少佐



バルツ
医学

産業革命に追いつき
富国強兵を成し遂げる

終戦後
(1945~)



エドワーズ・デミング



生産性の視点を導入し
焼け跡から立ち直る



若者へはデータ
リテラシー教育を
ただやればいいのか？

思考、表現の武器としての国語の刷新が第一

Not this

- 小説、随筆の書き手の理解、言いたいことの推測
- 感想文。感じたことの書き連ね。建設性のない批判
- 複雑な敬語。ソフトで角の立たない表現



But this

- 分析的、構造的に文章や話を理解し課題を洗い出す
- 論理的かつ建設的に物を考える
- 明確かつ力強く考えを伝える

“国語”

“コミュニケーション”

基礎と自信を身につける中等教育の質をいかに担保するか 、、、数学、英語、ICT

Newsweek

数学の「できない子」を強制的に生み出す日本の教育

2016年6月21日(火) 16時10分

舞田敏彦 (教育社会学者)

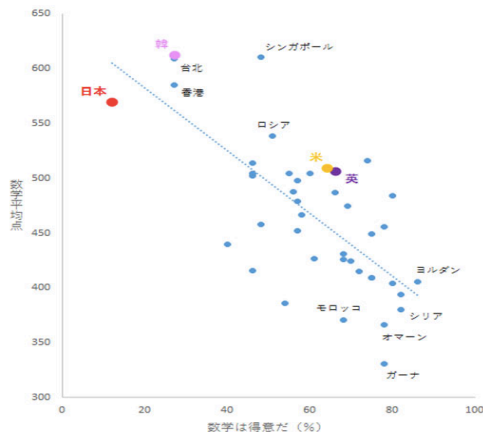
1,354

いいね!

1354

699

図1 数学の得意率と平均点 (中学校2年生)



* 横軸は、「とても当てはまる」+「少し当てはまる」の回答比率。
* IEA「TIMSS 2011」より筆者作成。

YAHOO! JAPAN ニュース

英語教員、TOEIC“合格”2割 京都府中学「資質」はOK?

京都新聞 2/10(金) 8:48配信

京都府教育委員会は9日、京都市を除く中学校の英語科教員で、本年度に英語能力試験TOEICを受験した74人のうち、府教委が目標として課した英検準1級に相当する730点以上を獲得したのは16人で、約2割にとどまることを明らかにした。最低点は280点で、500点未満も14人いたという。府教委は「英語科教員の資質が問われかねない厳しい状況だ」としている。

資料 : <http://headlines.yahoo.co.jp/hl?a=20170210-00000001-kyt-l26>
<http://www.newsweekjapan.jp/stories/world/2016/06/post-5365.php>



専門層・リーダー層 の育成について

このままでは米中と戦うことは非現実的に

PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE

Executive Office of the President
National Science and Technology Council
Committee on Technology

October 2016



2016年末に
ホワイト
ハウスから
相次いで
出された
AI白書

資料 : <http://japanese.joins.com/article/294/216294.html>
<http://www.nikkei.com/article/DGXLZO11372310W7A100C1MM8000/>
https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

① 中央日報

中国、人工知能開発に3年間で1000億人民元投入

2016年05月30日10時55分
[© 中央日報/中央日報日本語版]

コメント 0

中国が人工知能(AI)開発に1000億人民元(約1兆6800億円)程度を投じる。

日本経済新聞

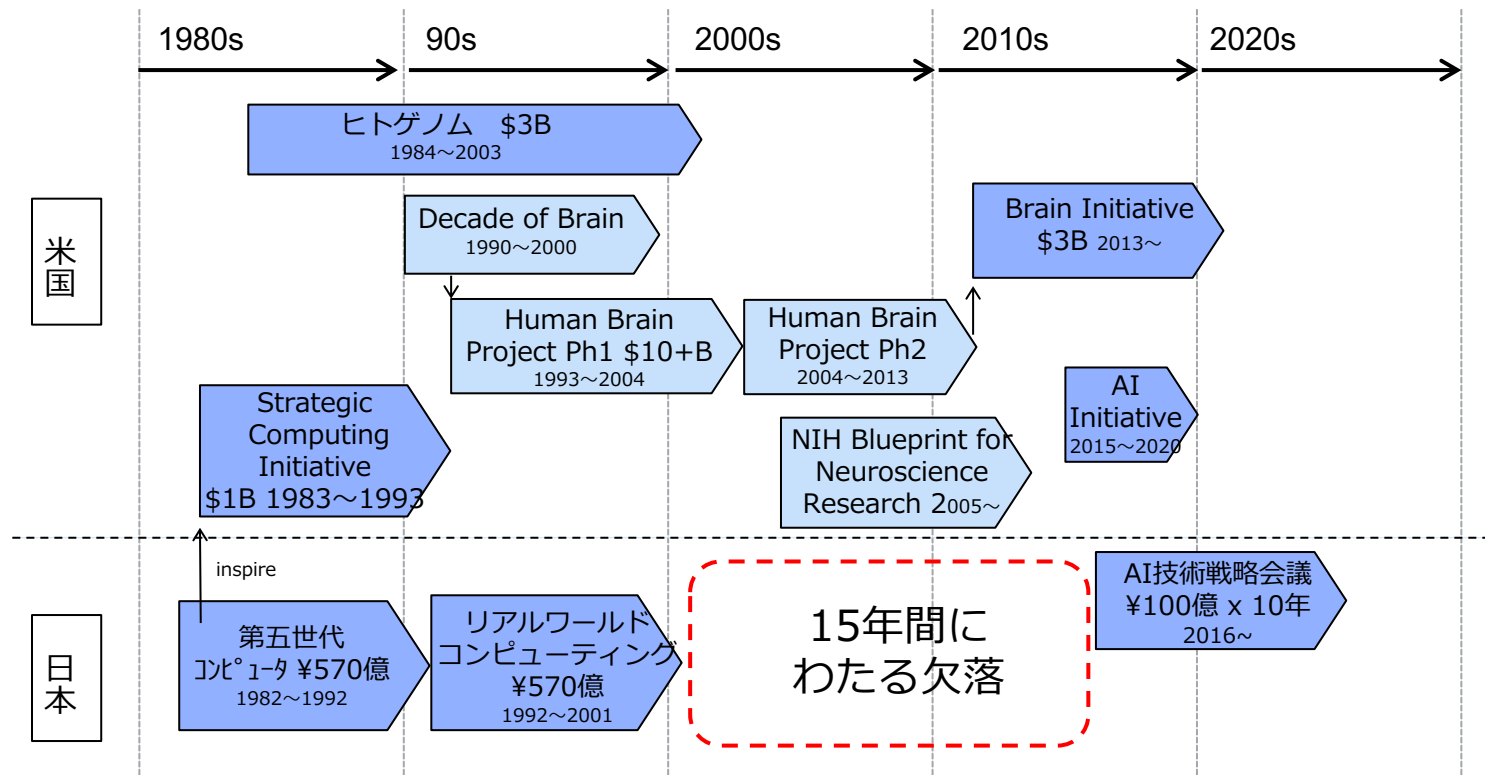
中国 次世代通信に5兆円 3社IoT先導狙う

2017/1/6 1:19 [有料会員限定]

【北京＝多部田俊輔】中国通信大手3社は2020年までに3千億元(約5兆円)規模を投じ、次世代の無線通信規格である第5世代(5G)の通信網を整備する。あらゆるモノがネットにつながる「IoT」や、自動運転のインフラとなる5Gが世界最大の中国市場でいち早く普及することで、中国発の技術やサービスが事実上の世界標準に近づきそうだ。

国力に見合ったグランドチャレンジを複数掲げるべき

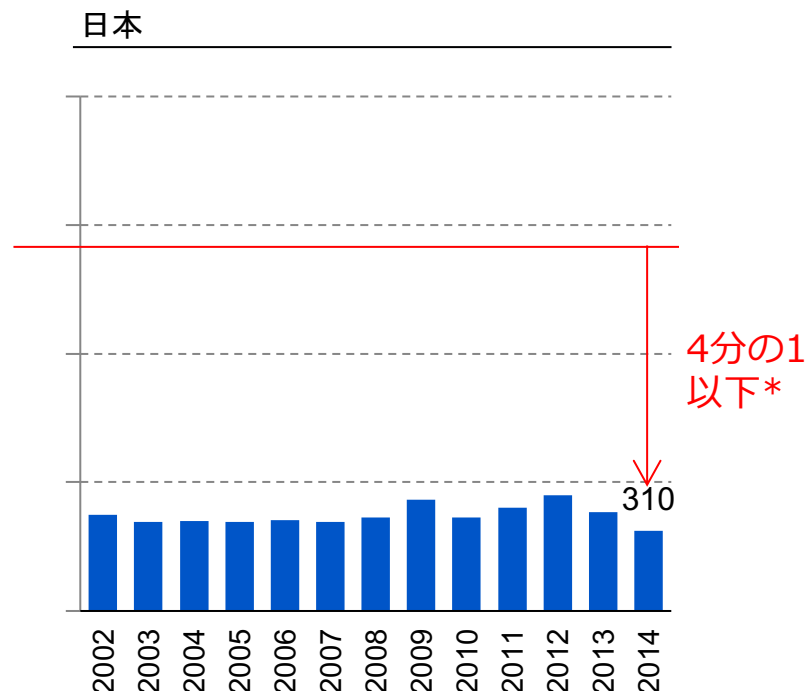
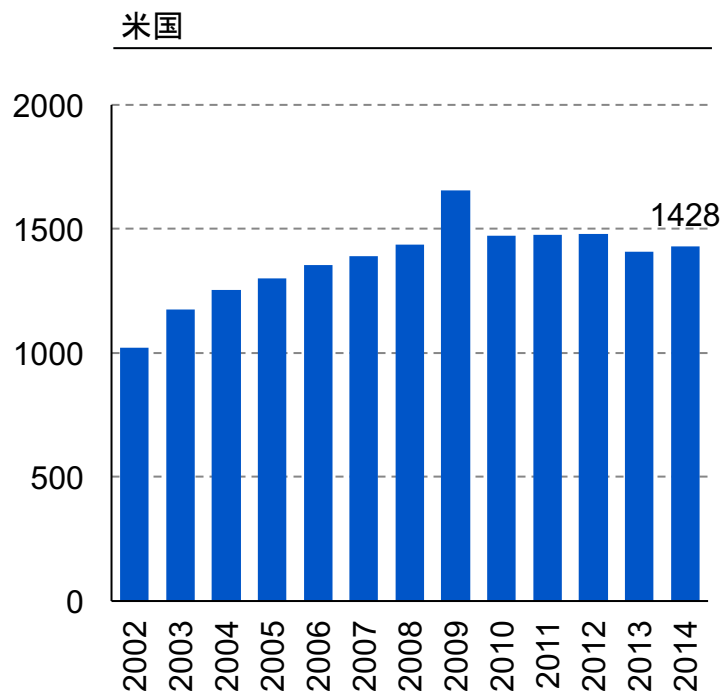
科学技術分野における日米の大型プロジェクト





高度人材育成のための原資強化も必須

政府の科学技術予算の日米比較 (億ドル)



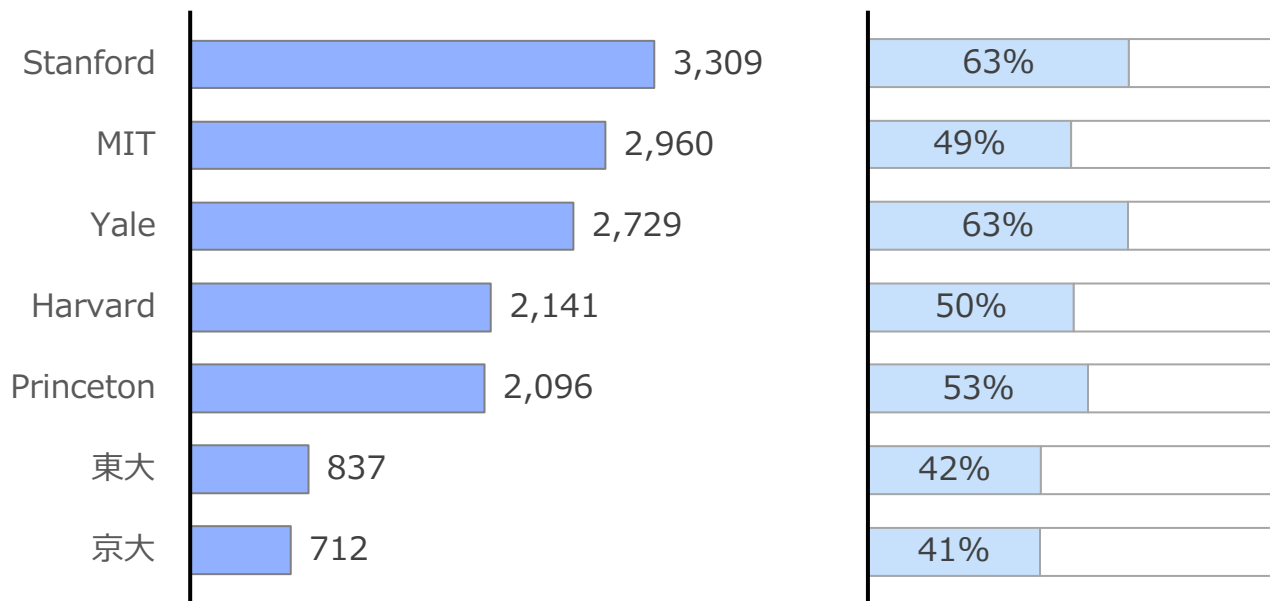
* 人口は約2.5分の1 (2013)

資料：米国：2014年度大統領予算教書における研究開発予算の概要、日本：文部科学省「科学技術予算に関する資料」、117円/\$で換算

日米の大学の資金力の差は大きい

大学の総支出*/学生
(100ドル/学生: 2015)

総支出に占める人件費率**
(%: 2015)

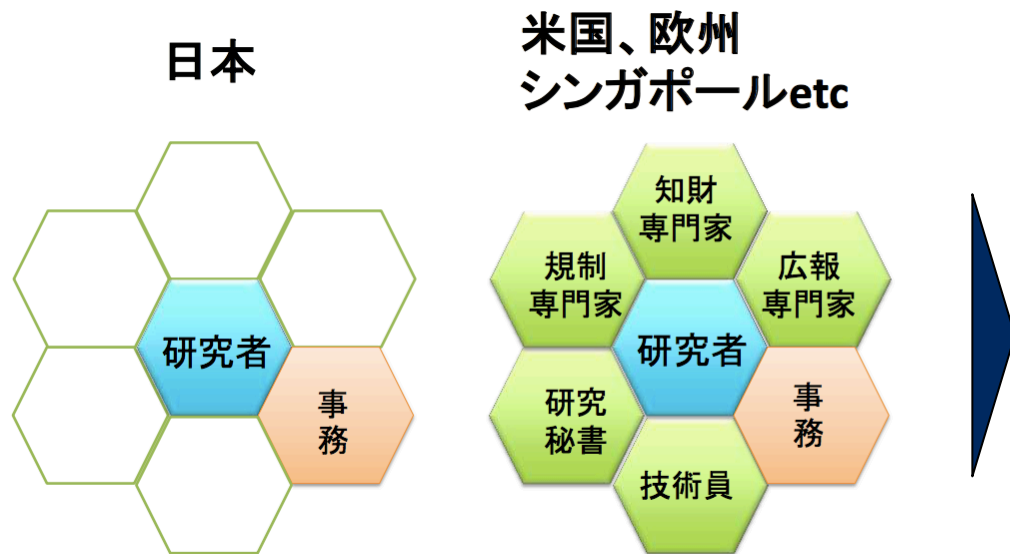


- 国際的競争力の
ない給与
- スタッフ不足
- リノベーション
されないビル群

* Operating expense (簡便のため\$1=¥100で換算) **給与に加えbenefit (諸手当) 含む
資料: 各校financial report、学生数 (学部、院のenrollment) に基づき安宅和人分析

この結果の一つが研究・教育環境の違い

研究開発人材の環境比較



研究者のコメント



「米国から日本に帰ってくると極端に雑務が多く、しかも政府の委員などの仕事が多く到底まとまった仕事ができない」
(分子生物学者)

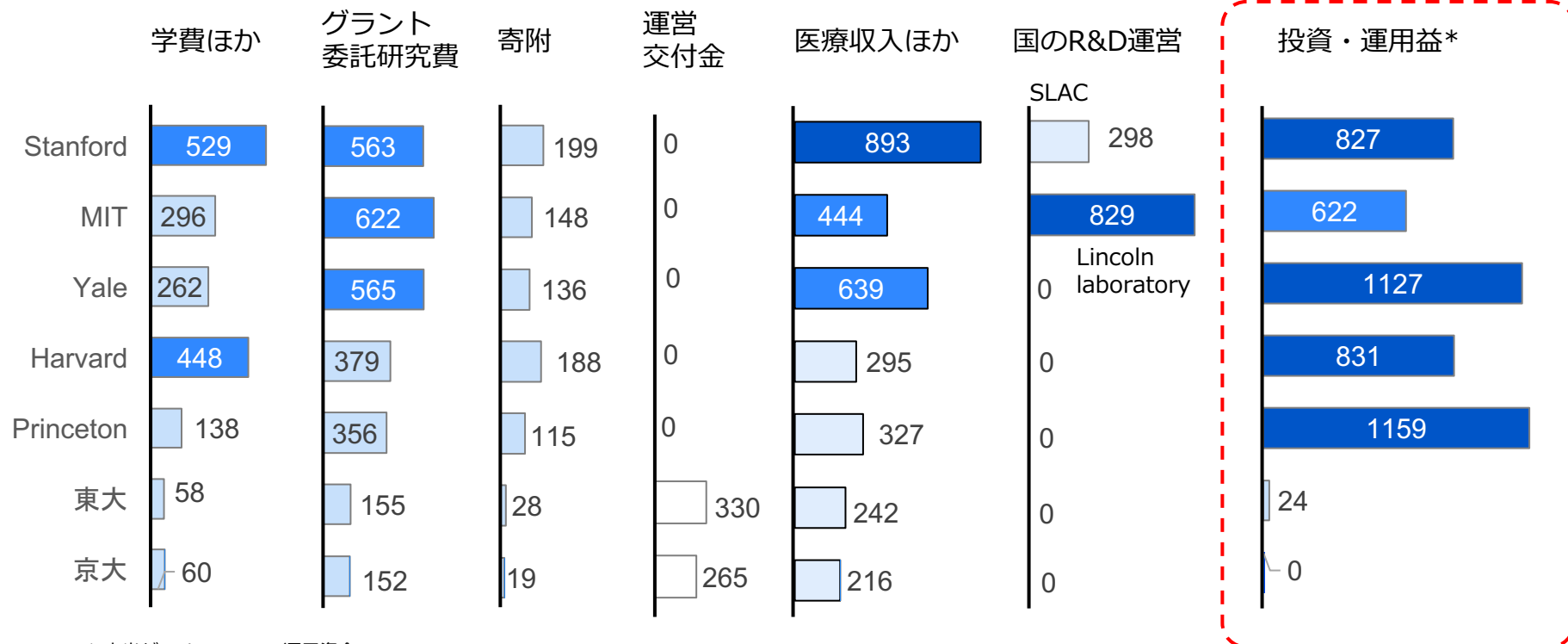


「シンガポールの大学のFacultyに入ると事務員だけでなく、大学スタッフにかなり専門的な知識と経験を持ったひと（中にはPh.D. holderも）がたくさんいる。この時期、入試にかかりきりの日本の大学では考えられない環境」
(情報系研究者)

最大のギャップは投資・運用益

大学別収入内訳/学生* (100ドル/学生: 2015)

■ > 800
■ > 400



* 大半がendowmentの運用資金

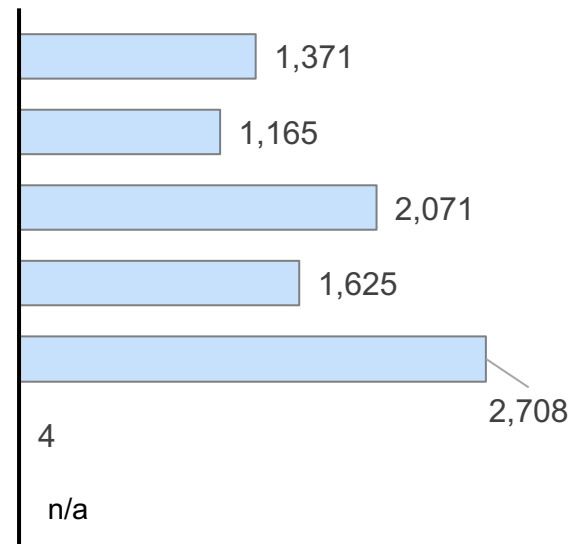
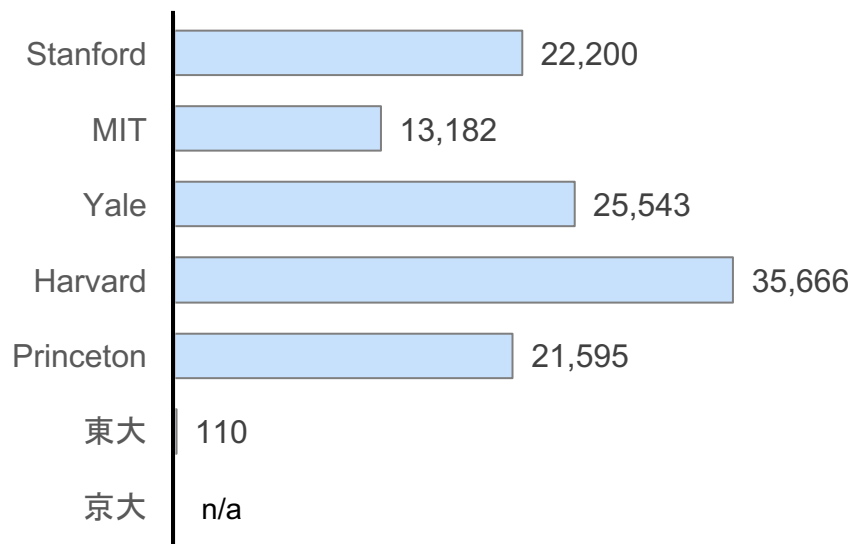
資料: 各校financial report、学生数 (学部、院のenrollment) に基づき安宅和人分析(簡便のため\$1=¥100で換算)



人材開発に向け国家的なendowmentを立ち上げるべき

大学別endowment総額比較
(\$ million : 2015)

学生一人あたりendowment比較
(千ドル/学生: 2015)



When
東大=1

340

289

513

403

671

1

n/a

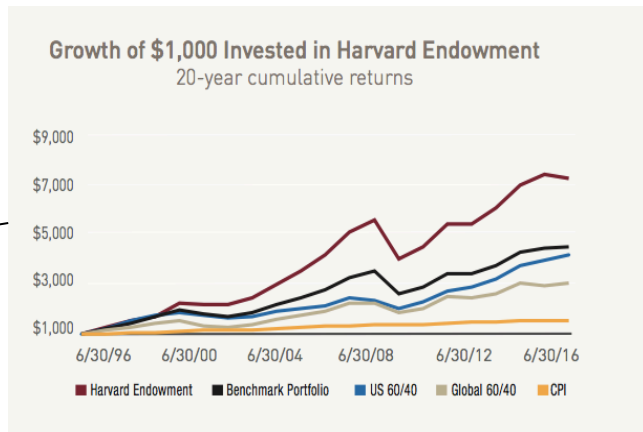
各校兆円単位の運用資金

学生1人億円単位



国力強化に向けた国家endowment構築（案）

1. トップ研究大学* (research university) 強化費用として**運用基金**として**10兆円**程度準備
2. 世界トップクラスの運用professionalを任命
、、、**平均7%以上**の運用益創出を目指す
3. 基金の3.5%程度（運用益の半分）を予算化。基本1/2を教員・サポートスタッフの人件費、1/4を人材育成グラント**、1/4は施設のリノベーションに
4. 大学などの教育研究機関***への寄付に対する免税措置 (education gift)
5. 企業が従業員や配偶者のこれらの寄附にマッチして寄附することも免税 (matching gift)



Giving to Yale

Matching Gifts

Do you or your spouse work for a company

Many employers sponsor matching gift programs. [Click here](#) to find out if your company has a matching gift program.

Not sure? [Find out here.](#)

* 希釈を避けるため10校以下が望ましい。世界的な研究力、優秀なPh.D.を生む能力、施設の老朽化などの必要性を鑑み指定（旧帝大+東工大+総研大など）

学費補助のスカラーシップ、生活費（stipend）支給など（米国同様、日本国民、永住権保持者を優先） * 幅広く対象

資料：Harvard University financial report FY16, <https://ces.commerce.yale.edu/givingtoyle/gifts.cgi>をもとに安宅和人試算

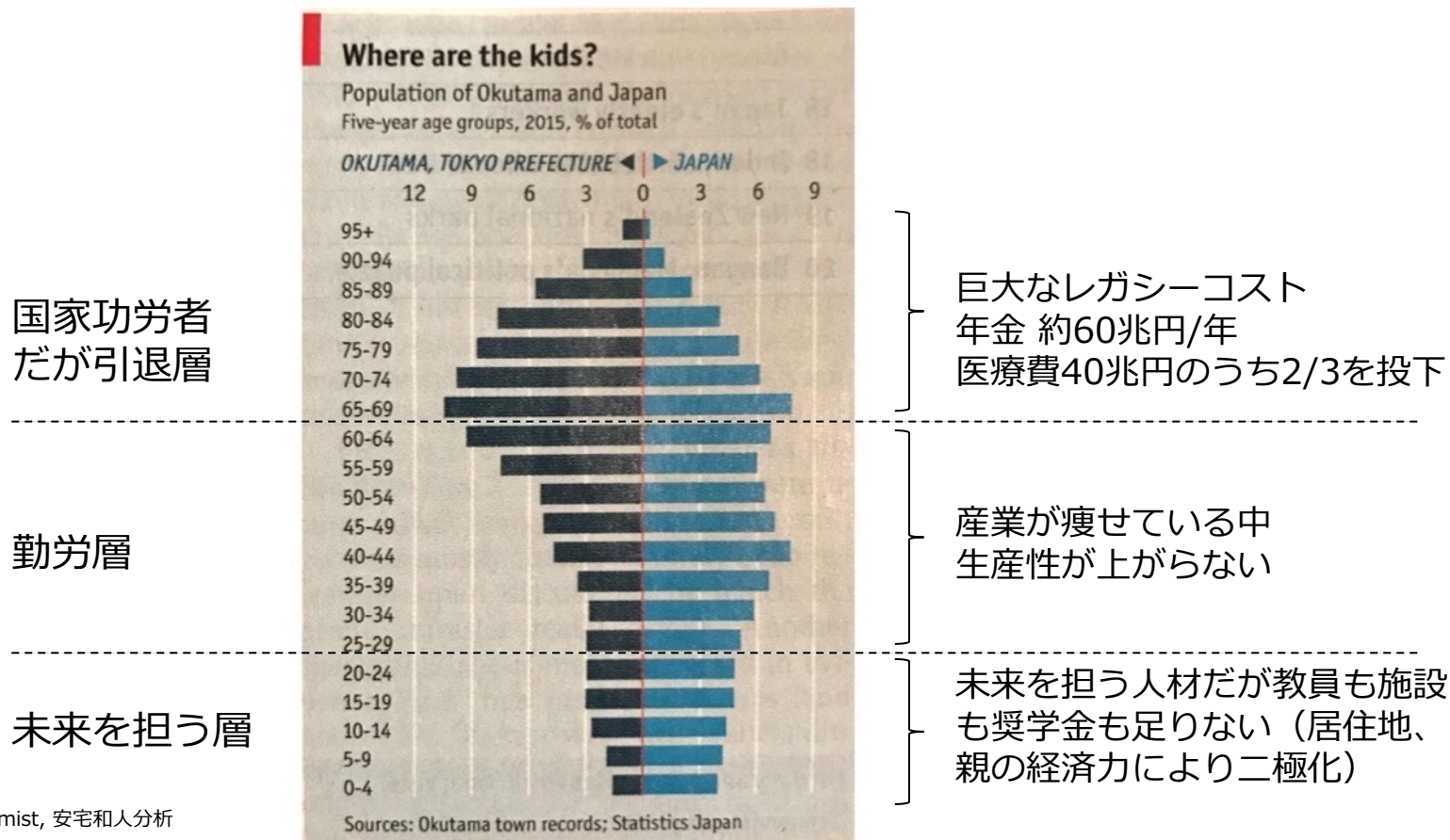
(参考) 米連邦政府の研究開発委託

Federally Funded Research and Development Center (FFRDC)



名称					
委託元	NASA	国防省	国防省	エネルギー省	エネルギー省
活動	NASAの無人探査機等の研究開発及び運用	空防に関する先端技術開発	ネットワークシステムの防御	エネルギー及び環境課題に質的変容をもたらす解を提供する	高エネルギー物理実験により宇宙の理解を再定義する
年間予算	\$1759M (2015)	\$956M (2016)	\$346M (2015-20平均)	\$785M (2014)	\$448M (2016)
運営	 Caltech				 STANFORD UNIVERSITY

国家の経営としてのリソース最適化を検討すべき



まとめ



1. 国富を生む方程式が変化、、、妄想力とそれを形にする力がカギ
2. 日本は情報産業革命の第1フェーズで大敗、、、人材課題は多面的に深刻。勝負は第2、第3フェーズ
3. データの持つ力を解き放った上で、見る力・決める力・伝える力が重要になる、、、これまでとは似て非なるdata professionalが必要
4. 3層+2で育てるべき
 - ① 若者に武器を、、、データリテラシーとアントレプレナーシップ
 - ② 境界・応用領域を含む専門家層とリーダー層は必須
 - ③ ICTエンジニアとミドル/マネジメントを再生
5. 千載一遇のチャンスを活かし世界の才能を取り込む
6. 基盤となる思考、表現の武器としての国語の刷新、人材開発に向けた国家的なendowmentを立ち上げるべき
7. 以上の実現に向け、国家全体のリソースの最適化を検討すべき