

世界大学ランキングの“恣意性”について

On “Arbitrariness” of World University Rankings

調 麻佐志（東京工業大学）
SHIRABE, Masashi (Tokyo Tech)

恣意性（大辞林）

Definition of *arbitrariness* (Oxford dictionary)

- 1. The quality of being based on random choice or personal whim, rather than any reason or system.**
 - その時々 of 気ままな思いつき。自分勝手な考え。
 - 物事 of 関係が偶然的であること。
- 2. A lack of restraint in the use of authority; autocracy.**

Big leaps of Turkish universities in THE rankings

2013-14

2014-15



Higgs粒子論文が原因

Attributed to Higgs boson papers

ATLAS collaboration

Ankara Univ
Dumlupinar Univ
Gazi Univ
TOBB Univ Econ & Technol
Boğaziçi Univ
Dogus Univ
Gaziantep Univ
Istanbul Tech Univ
Middle E Tech Univ

CMS collaboration

Cukurova Univ
Middle E Tech Univ
Boğaziçi Univ
Istanbul Tech Univ

Boğaziçi Univ. (as of Autumn 2015)

Total citations from papers published in 2012

: 12362

Those from Higgs boson papers

: 2552+2409 →40%

ランキング機関の対応

Responses of THE & QS

• THE

- 【kilo-authors論文の排除】 This year [2015] we have removed the very small number of papers (649) with more than 1,000 authors from the citations indicator.
- 【変形分数カウントの導入】 [A] new fractional counting approach that ensures that all universities where academics are authors of these papers will receive at least 5 per cent of the value of the paper, and where those that provide the most contributors to the paper receive a proportionately larger contribution. (2016)

• QS

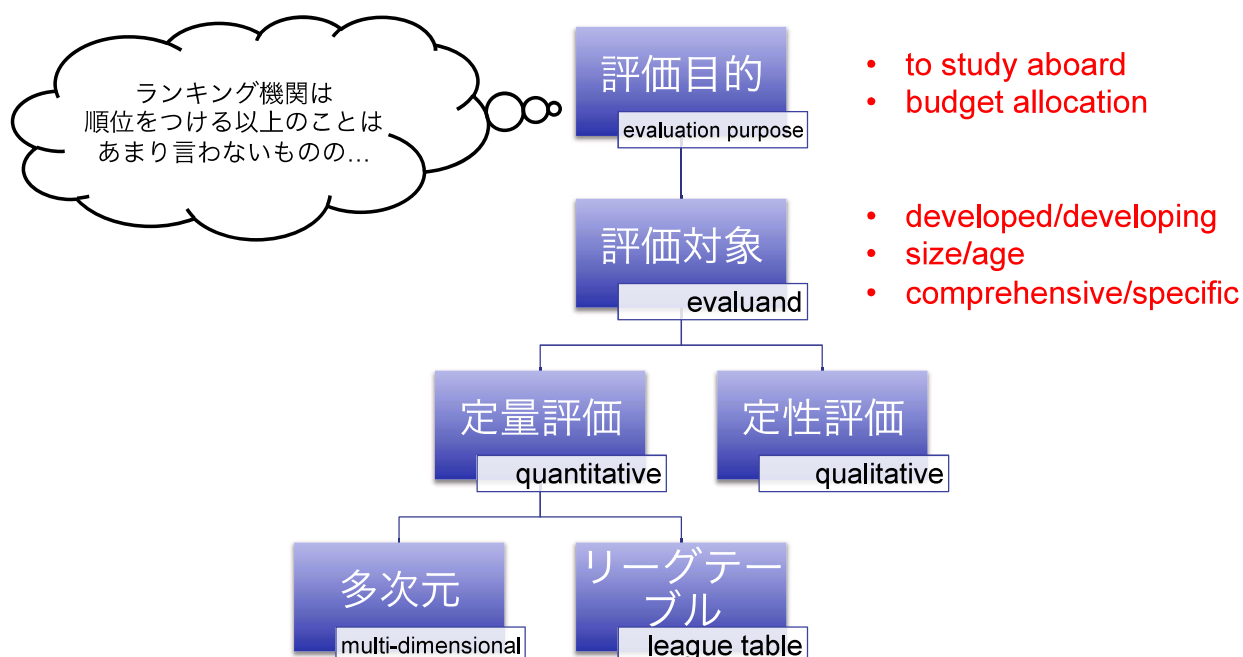
- 【10機関論文の排除】 [w]e are proposing to exclude papers with more than 10 affiliated institutions – representing 0.34% of records. (2015)

典型的なアドホックな決定 (ad hoc decisions)

なぜ千人/10機関？ 普通の分数カウントは？

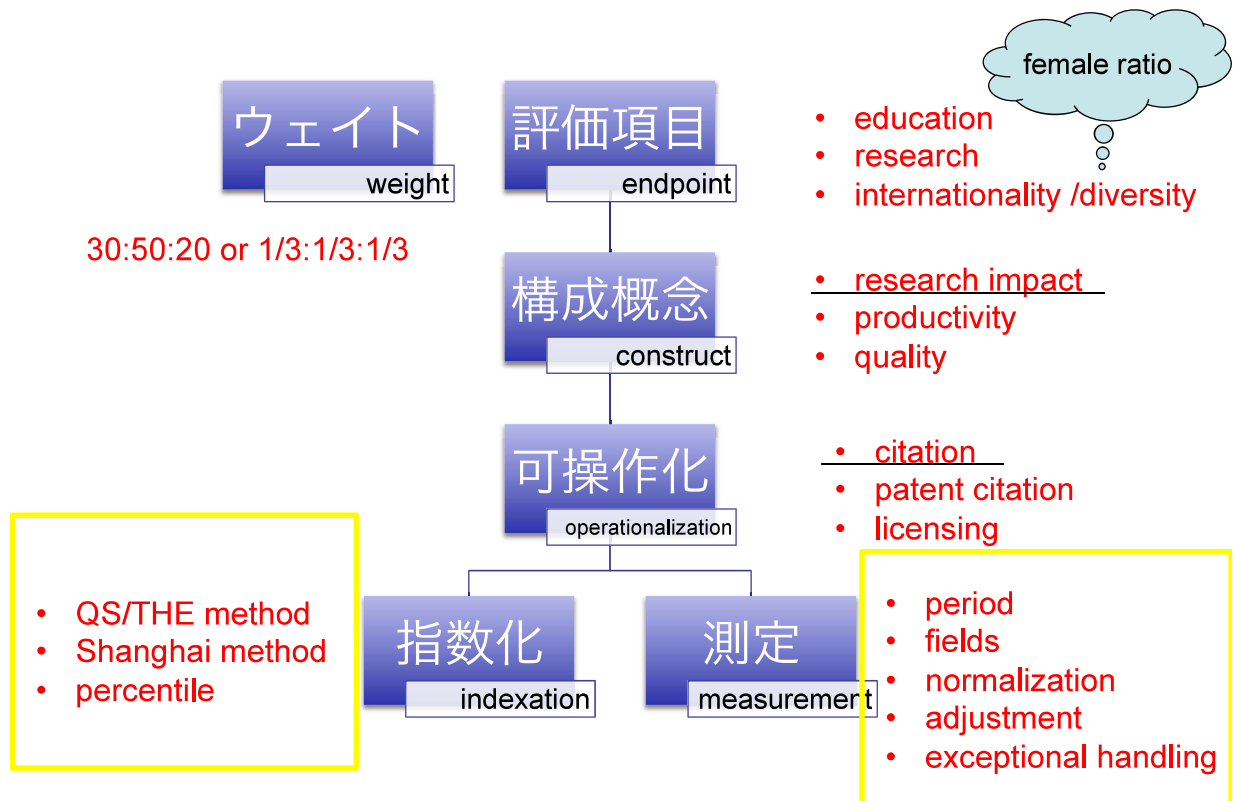
ランキングに“恣意性”は付き物

Rankings are not free from arbitrariness.



ランキングに“恣意性”は付き物 2

Rankings are not free from arbitrariness.



恣意性の次元 (試論)

“Dimensions” of arbitrariness

- **内在的 (intrinsic) 恣意性** → ウェイト設定
- **適合性/適用可能性 (relevance)**
 - 卒業生ノーベル賞・フィールズ賞受賞者 (ARWU) → 教育
 - 「パフォーマンス」のパタンにかかるもの (偶然的 accidental)
 - THEのcitations
 - 国際性 (留学生比率が高いほど高スコア)
 - 「パフォーマンス」外の要因・特性にかかるもの (系統的 systematic)
 - 国補正 (country adjustment) → 非英語雑誌 (non-English J.)の扱い?
 - 中心-周縁 (center and periphery)
 - reputationにおける総合大学と単科大学
 - 国際性 (EU、英語圏、...)
- **信頼性 (reliability)**
 - 「操作」に対する頑健性 (robustness) → 自己引用 / 過大申告
 - 安定性 (stability) → methodologyの変更は...
 - 一貫性 (consistency) → 外部基準 (QS vs. THE)、ウェイトには?
- **効率性 (efficiency)** → **他の次元とトレードオフ**
 - 教員一人あたりの学生数 = 教育 ?

THEランキングのCitation項目

Citation in THE rankings



• Citations (research influence)

- 文献一本あたりの引用数(文献種類／分野等で正規化)
normalized citations / document type... (FWCI)
- 思考実験(thought experiment)
 - Institute A: 5, 2, 2, 2, 2, 2 → 1.5
 - Institute B: 100, 0, 0, 0 → 1.1

低引用論文の出版は罰せられる。

Publication of low citation papers is “penalized”.

- THE社は「引用は教育効果もと...」

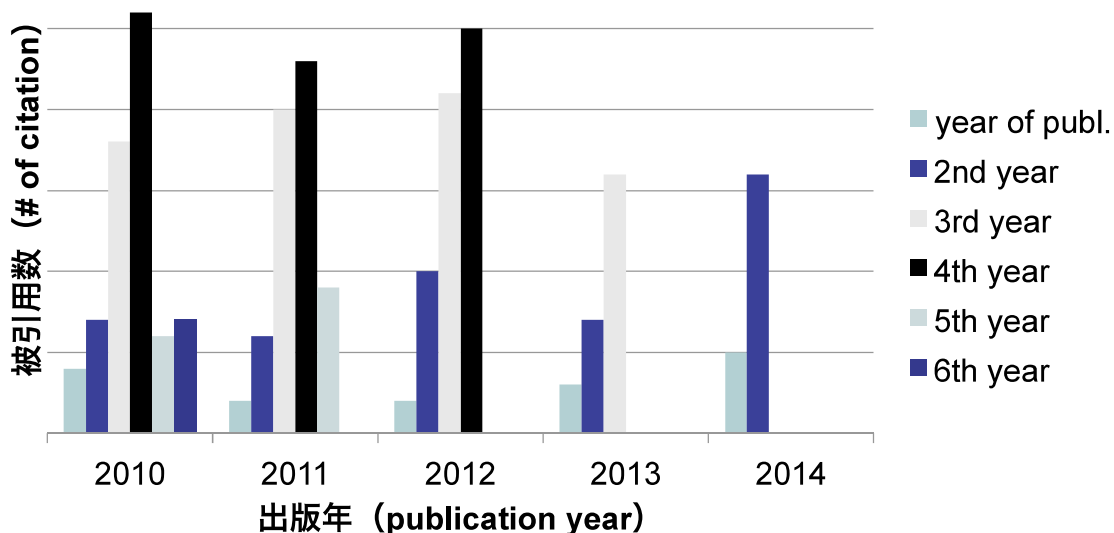
QSランキングで行う「教員一人あたり引用数」
が妥当であるという主張ではないことに注意。

正規化した被引用数5年平均は系統誤差を？

Does a five-year average of normalized citations invite systematic errors?



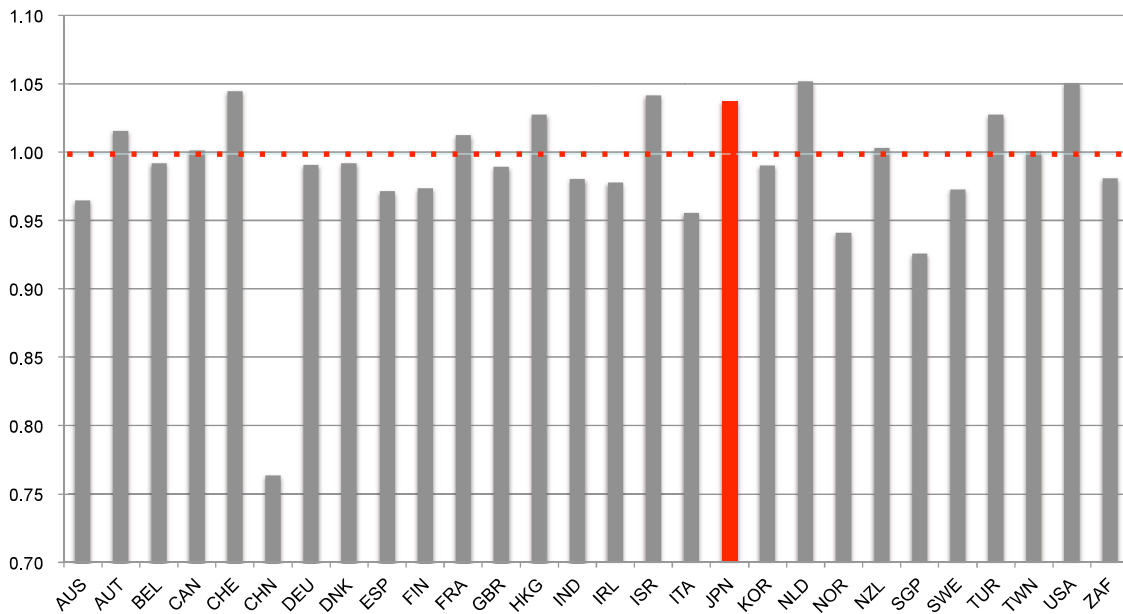
THEランキング2016では2010～14年出版論文の
2010～15年の正規化被引用数 (FWCI) を集計



トップ10%論文比率の比 (top 10% 2010/ top 10% 2013)
Ratio of shares of top 10% cited papers (2010/2013)



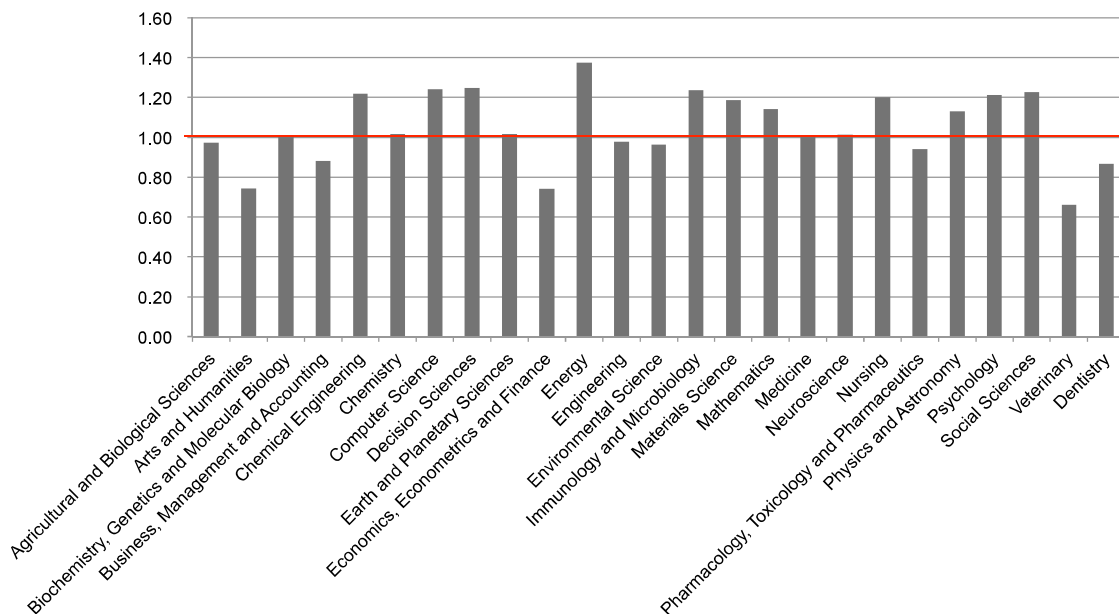
(by country)



日本のトップ10%論文比率の比 (2010/2013)
Ratio of shares of top 10% cited papers (2010/2013)



分野別(by field)



サンプリングのスコアへの影響

Effect of sampling on score



• 素データのスコア化

→FWCIで確認

Transforming raw data into scores

– QS & THE ranking method

- 平均 μ と標準偏差 σ を求め、当該データが $N(\mu, \sigma)$ に従うとしたときの下からのパーセンタイルをスコアとする。

– ARWU ranking method

- 最上位データを100とし、以下は比例した値がスコア。
→素データと本質的には同じ。

– それ以外の方法 (other possible methods)

- 素データのパーセンタイル
- 正規分布でフィッティング→パーセンタイル

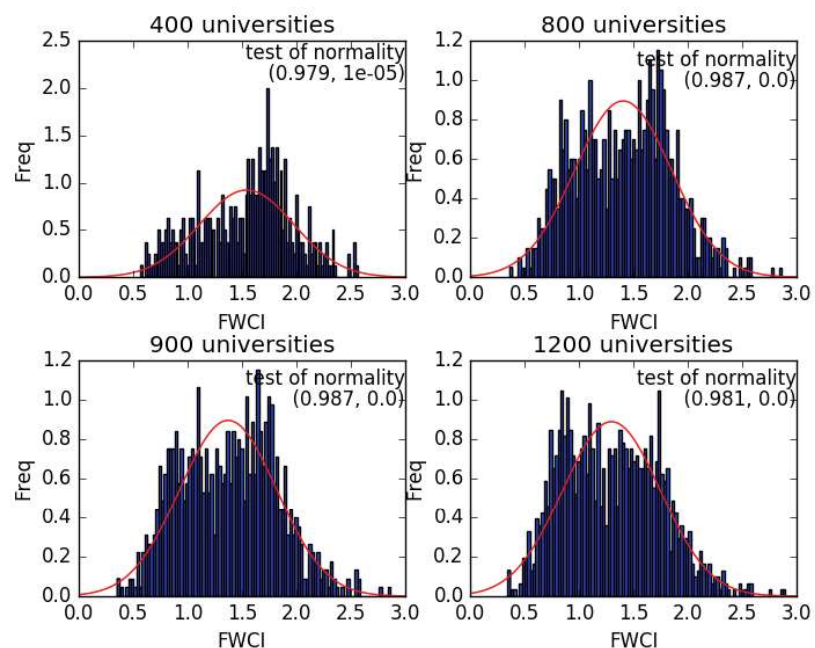
そもそもFWCIの機関平均は正規分布？

Are averages by inst. of FWCI normally distributed?



• 論文数上位

400(old THE)、
800(new THE)、
900(QS)、
1200(ARWU)大学のFWCIの分布に対してシャピロ-ウィルク検定 (the Shapiro-Wilk test)



正規分布でない。

(フィッティングによるスコア化は×)

sampling indexing method	400 largest univ.			800 largest univ.			1,200 largest univ.		
	THE/QS	ARWU	%	THE/QS	ARWU	%	THE/QS	ARWU	%
Harvard	98.5	96.7	98.7	99.2	86.7	98.9	99.6	85.6	99.2
MIT	99.7	87.7	99.4	99.4	88.7	99.3	99.7	87.7	99.4
Stanford	99.1	100.0	100.0	99.5	89.7	99.6	99.8	88.6	99.6
Caltech	98.9	98.8	99.2	99.4	88.6	99.2	99.7	87.6	99.4
Tokyo	31.8	52.2	31.0	44.2	46.8	44.4	53.7	46.2	54.0
Kyoto	24.5	48.5	26.5	36.1	43.5	38.4	45.4	43.0	47.7
Osaka	17.4	44.3	21.3	27.5	39.8	32.0	36.1	39.3	41.2
Tohoku	22.1	47.2	25.3	33.3	42.4	36.4	42.5	41.8	45.8
Tokyo Tech	20.8	46.5	24.3	31.7	41.7	35.4	40.8	41.2	44.8
Kobe	14.1	42.0	17.5	23.2	37.7	27.1	31.3	37.2	36.4
NUS	72.7	70.3	71.9	81.5	63.1	81.4	87.1	62.3	86.7
Peking	36.0	54.1	32.8	48.6	48.5	47.3	58.1	48.0	57.2
SNU	31.6	52.1	30.8	44.0	46.7	44.3	53.6	46.2	53.8
NTW	28.5	50.6	28.5	40.6	45.4	41.9	50.1	44.8	51.1

まとめと今後の展開 Summary and future plan

- ランキング作成のすべての段階に恣意性がみられる。その恣意性は結果に大きな影響を与える。

There can be seen arbitrariness in all the processes to produce world university rankings. Such arbitrariness affects ranking results heavily.

- ランキングを作成する以上、恣意的判断は不可避である。多少でもマシなものを選ぶことを支援するような研究が必要。

As we cannot avoid arbitrary decisions when producing world university rankings, research to assist people's selection of least-worst options is required.