The European Commission's science and knowledge service

Joint Research Centre

European Commission

Step 1 & 2
Frameworks and Indicators

Dániel Vértesy

COIN 2018 - 16th JRC Annual Training on Composite Indicators & Scoreboards
05-07/11/2018, Ispra (IT)

European Commission

European Commission

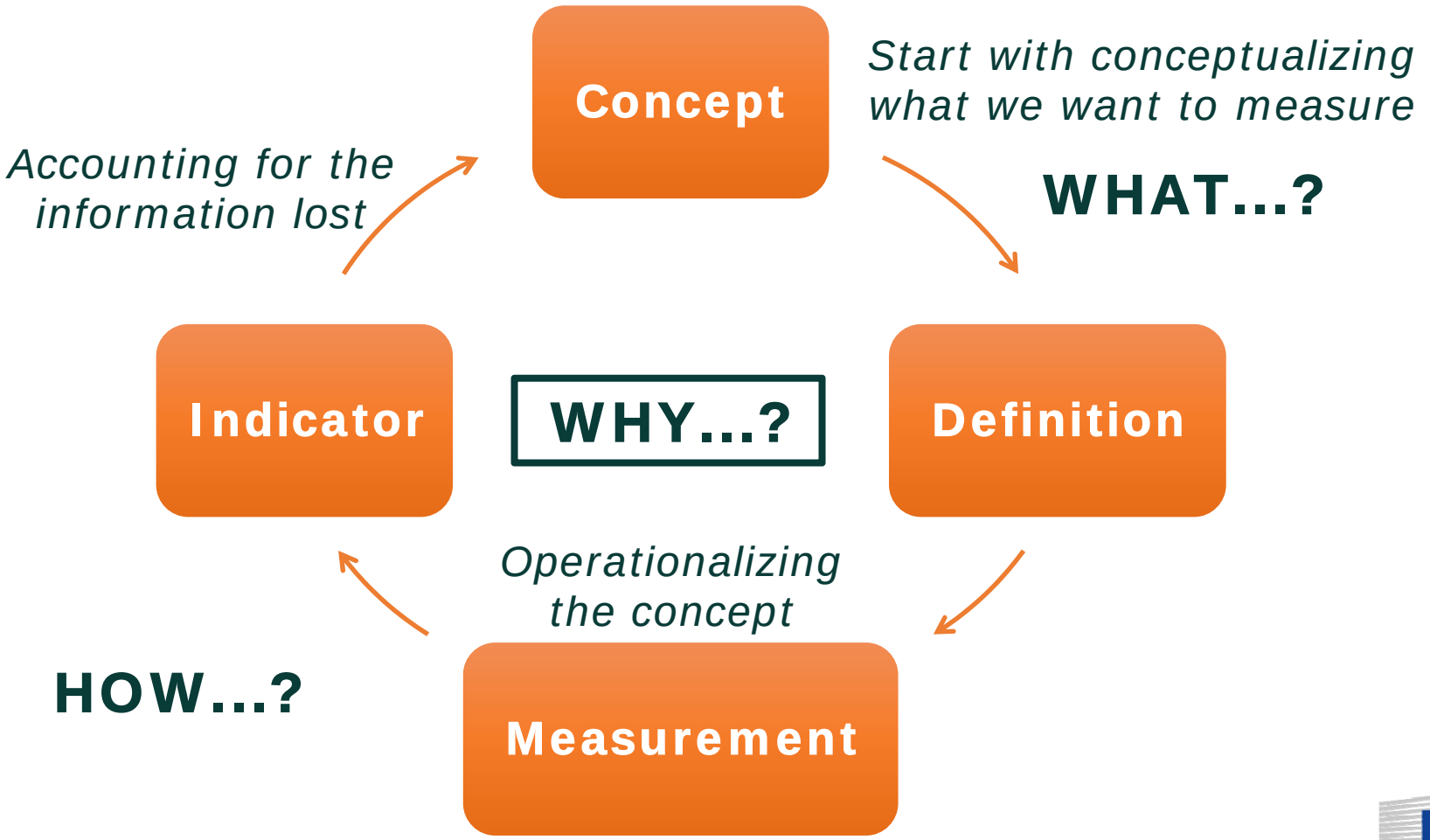
3 questions to clarify

? **What** do we want to measure?

? **Why** do we want to measure it?

? **How** do we want to measure it?

Operationalizing the Concept



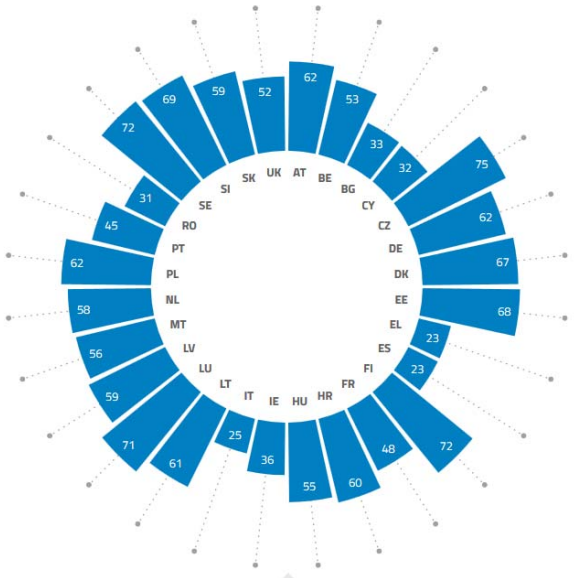
Measure the Skills System performance

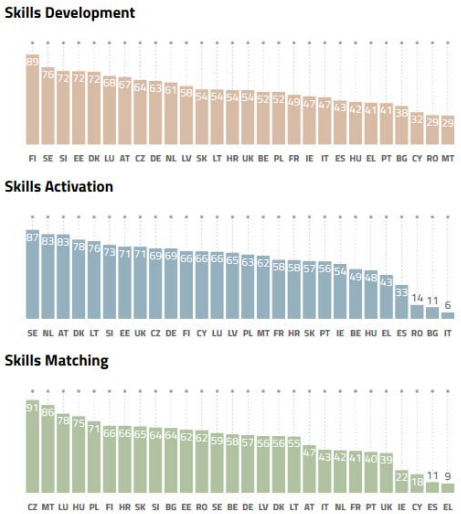
CEDEFOP European Centre for the Development of Vocational Training

- Provide evidence for European vocational education and training (VET) policy (trends; challenges)
- forecast the demand and supply of skills

The **European Skills Index** (ESI) is Cedefop's composite indicator measuring the performance of EU skills systems.

- A country's ~ delivers enhanced skills to the population through compulsory education or post-compulsory education and training;
- includes formal & informal E&T





European Commission

What is the role of a Theoretical Framework?

- **Guides** the choices for…
  - pillars (or dimensions);
  - weights;
  - aggregation methods;
  - etc.
- Supports the **interpretation** of results
- Command "stakeholder respect"

In sum: help answer the What/Why/How questions

European Commission

3 questions to clarify

What do we want to measure?

- SKILLS SYSTEM: multi-dimensional in nature; no obvious single indicator

Why do we want to measure it?

- A macro-level, comparative assessment of the skill system of Member States
- Originally: input to the autumn 2015 re-launch of the EUSP website
- Advocacy: naming (MSWI → ESI)


Measuring Performance of EU Member States Skills Systems and Prioritising Skill Mismatches

Measuring and comparing the relative situation and performance of EU Member States' skills systems

Deliverable FINAL.1 submitted to Cedefop

May 2015

How do we want to measure it?

European Commission

? HOW? – Raise guiding questions

- Why is a country's *skills system* important, what roles does it fulfill?
- What are the differences vis-à-vis closely related, established concepts? (= value added of a new index?)
  - i.e., Programme for the International Assessment of Adult Competencies (PIAAC)
  - What is the difference between *skills* and *employment* systems?
- Skills system vs. human capital?
- Links between individual and societal levels?
- Focus on Persons (employees) vs. Businesses (employers)?
- What comprises 'good' performance, and how can it be measured?
- What kind of data to use? (consistent, internationally comparable)
- What countries & years to cover?

The ESI Theoretical Framework

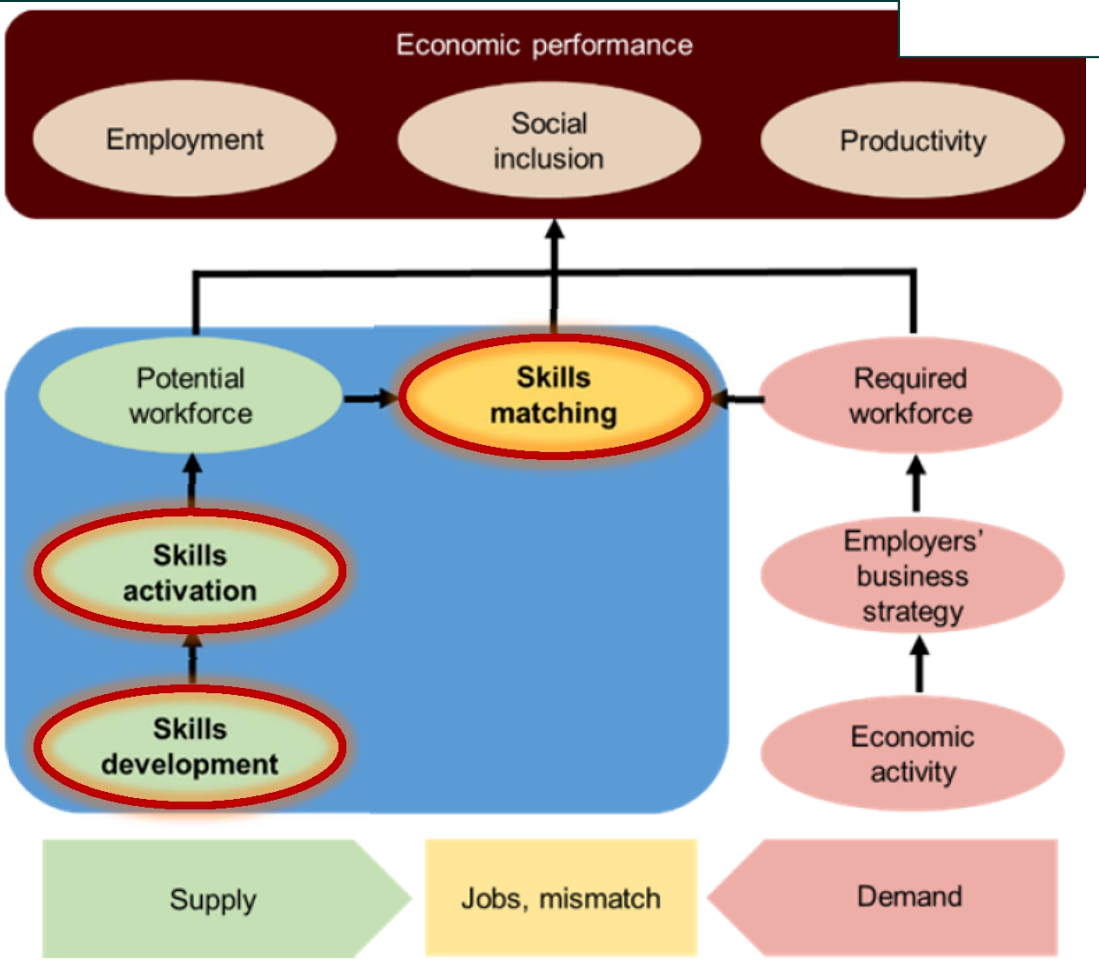
A county's skills system fulfills 3 Roles :

1. **providing an initial foundation** upon which individuals can continue to develop their skills

2. **delivering the skills** the country needs and/ or is anticipated to need in the future (including re-skilling and up-skilling);

3. **matching**, as far as possible, individuals' aspirations, interests, and abilities to the needs of employers.

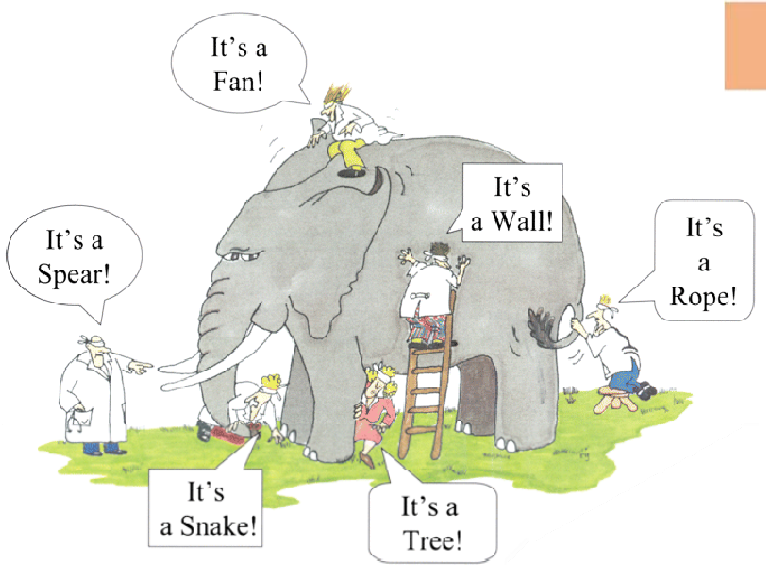
➔ **3 Pillars**

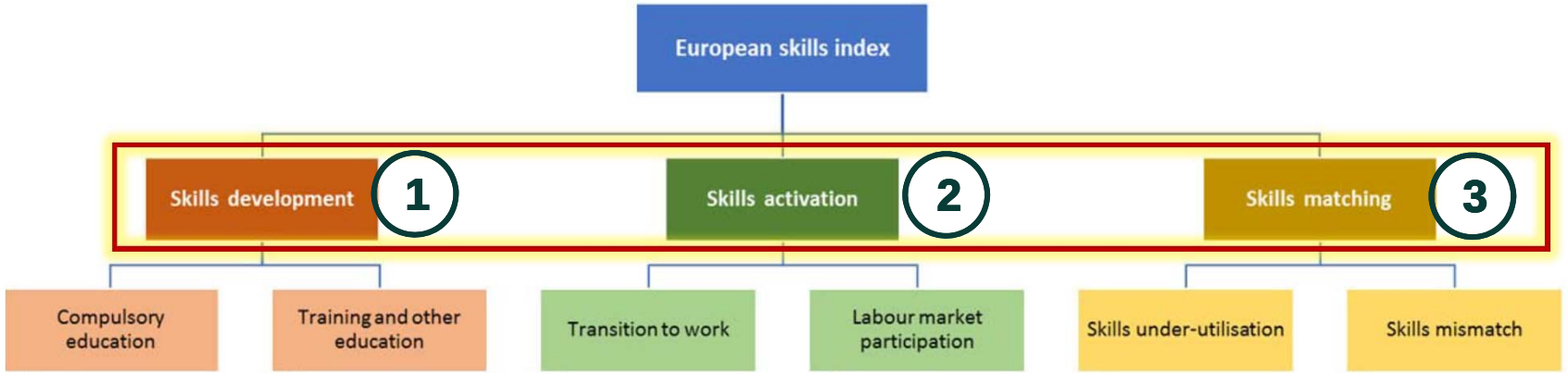


European Commission

The ESI Indicator Framework

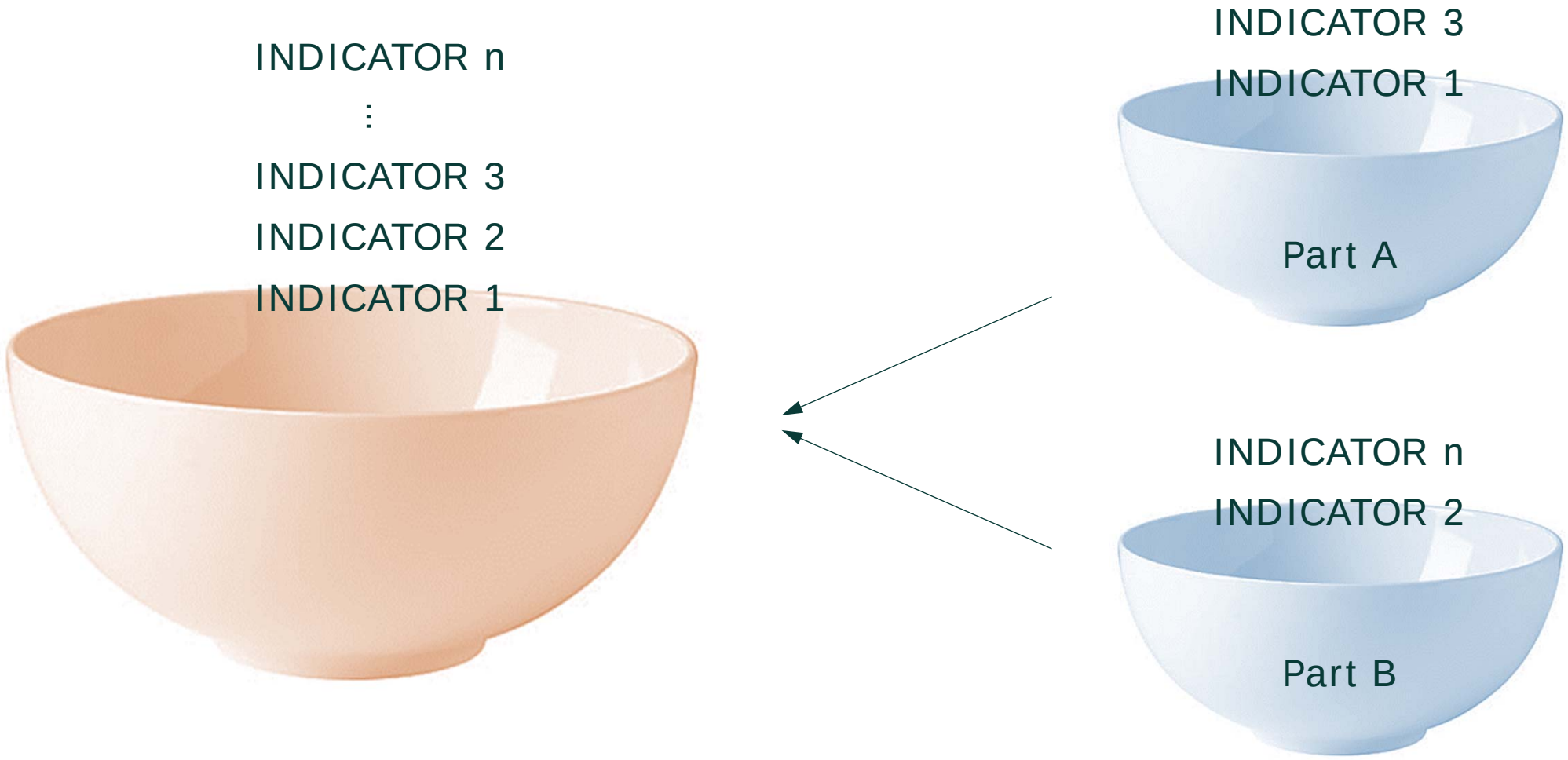
capture different aspects of a phenomenon



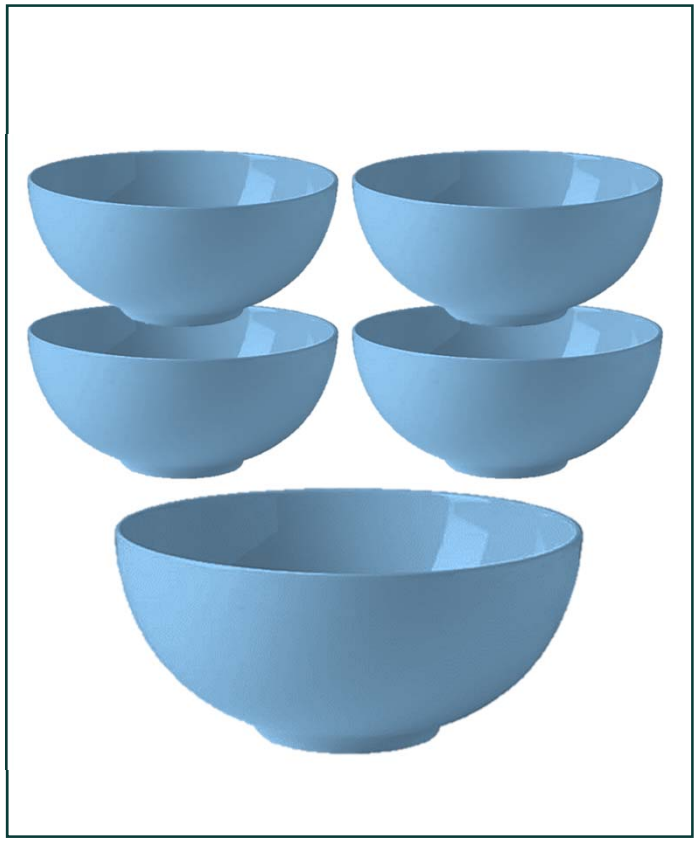


European Commission

How to fill a framework with indicators?



Frameworks



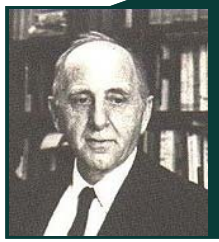
Is the quality of our life improving?

? What defines the “quality of our life”?

? What do we consider “improvement” or “development”?

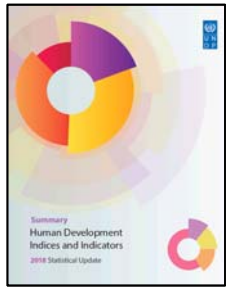
The Concept behind the Framework

GDP



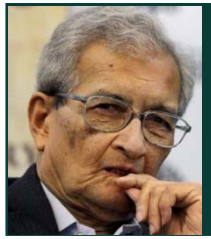
Measure wealth (production) *single measure* (S. Kuznets)

Human Development



Human Development Index (UNDP)

- 4 indicators,
- 3 dimensions



Capabilities and choices: *no single indicator!* (A. Sen)

Well-being

Better Life Index (OECD)

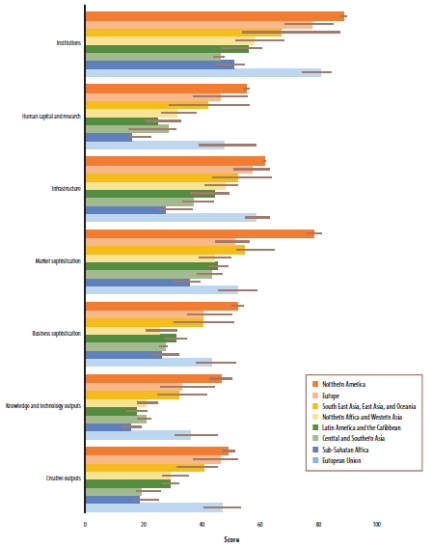
- 50 indicators,
- 11 dimensions

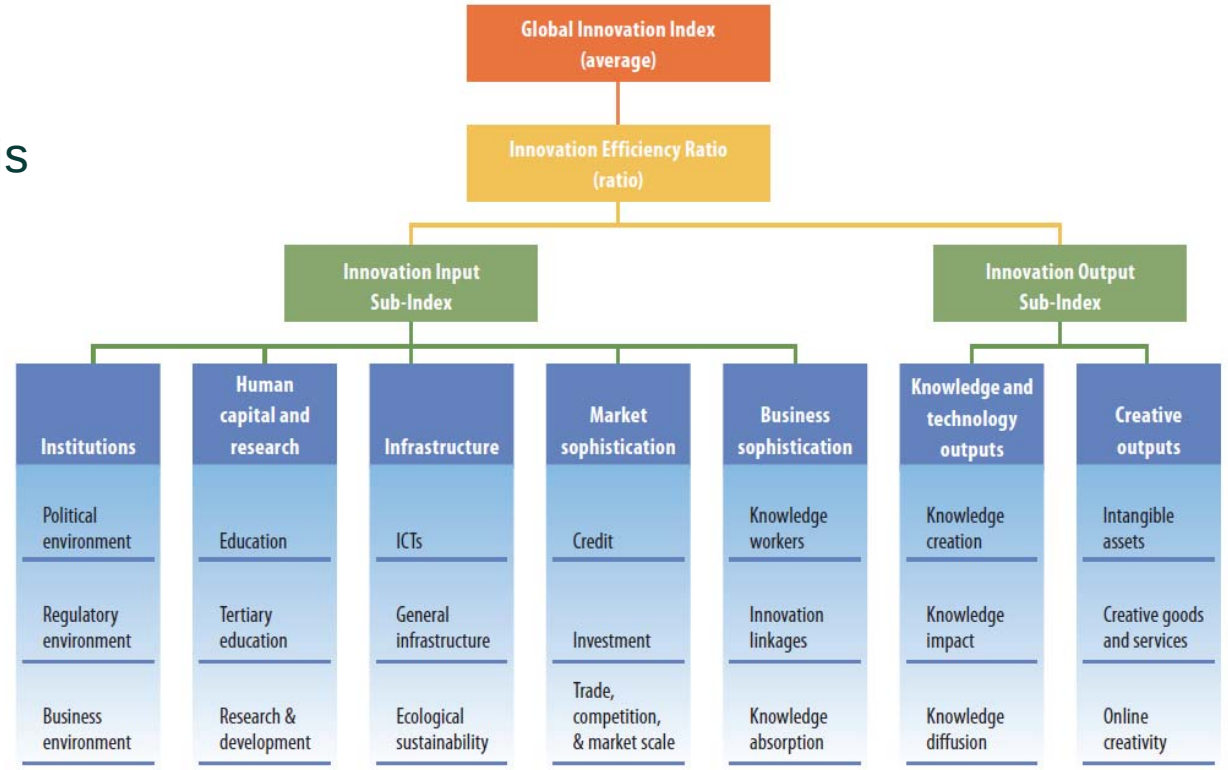
"There is more to life than the cold numbers of GDP and economic statistics"

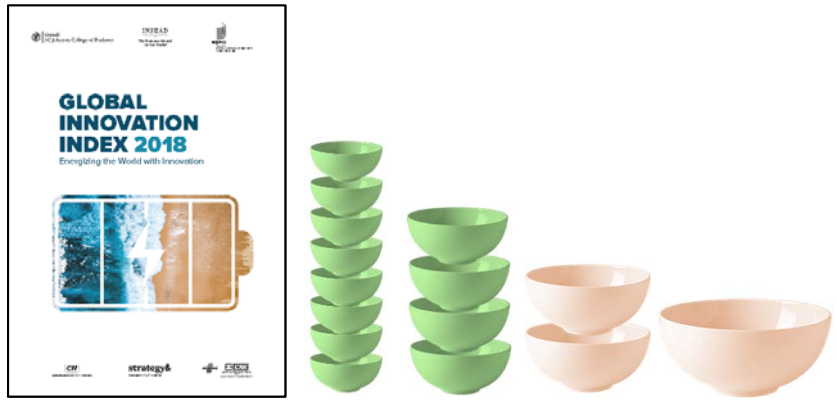
European Commission

Pillars & Sub-pillars: analytical building blocs

Pillar-level analysis







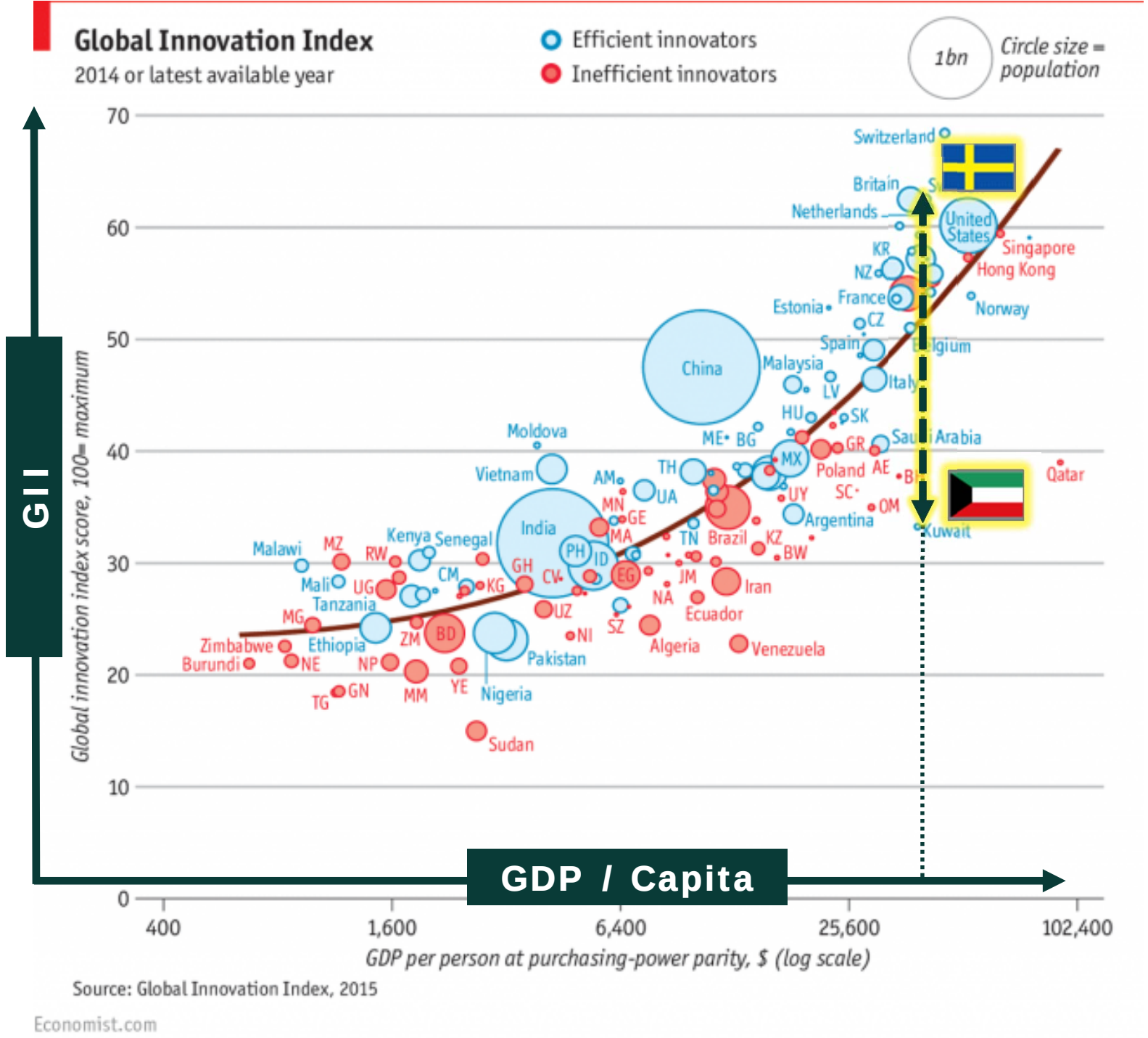
Global Innovation Index (WIPO, Cornell, INSEAD)

- 1 index
- 2 sub-indices
- 7 pillars
- 21 sub-pillars
- 82 indicators

European Commission

What is the value added of aggregate indices?

i.e., what GDP does not show



In sum...

Developing a Conceptual Framework

Developing a Theoretical Framework

- ✓ Have a clear definition of the concept
- ✓ Identify the sub-groups of the multi-dimensional concept
- ✓ Set up the selection criteria for underlying indicators
- ✓ Take the time to document your choices…

Selecting Indicators

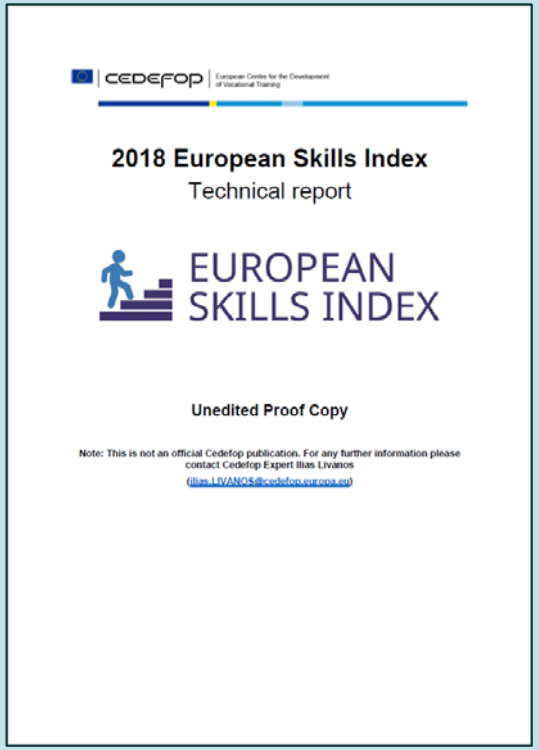
Selecting the ESI indicators

Criteria:

- Analytical soundness
- Measurability
- Country coverage
- Relevance

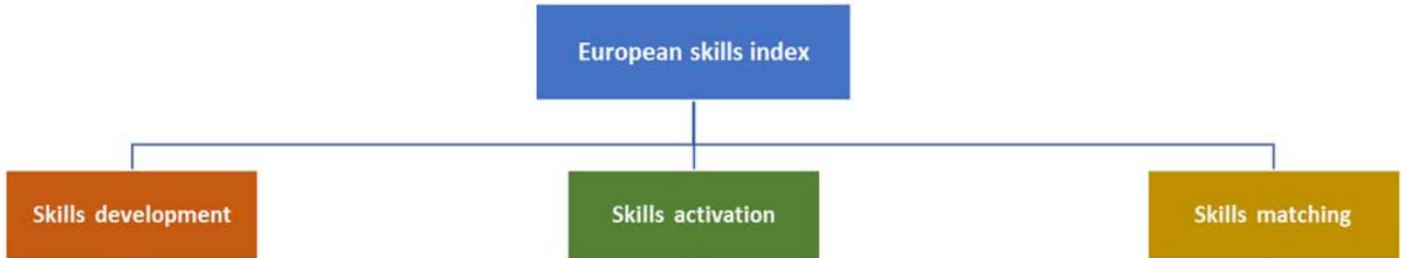
Iterative process

- Draft → Expert feedback → Meeting criteria...
- i.e.: discard ambiguous ‘expenditure indicators’



European Commission

The ESI Indicator Framework



Selecting the ESI indicators

Justifying the inclusion or exclusion of indicators

Table 1. The European Skills Index: Conceptual framework (right) and earlier working version (left).

Making Skills Work Index version 2017

Pillar (P)	Sub-pillar (SP)	Indicator group (IG)	Indicator (ind)	
1 P1: Skills Development	SP1: Compulsory education	IG1: Participation to compulsory education	Pre-primary participation	ind.01
			Upper secondary participation (aged 15-17)	ind.02
		IG2: Attainment from compulsory education	Upper secondary attainment (aged 15-64)	ind.03
			Reading, maths & science scores (aged 15)	ind.04
	SP2: Post-compulsary education and training	IG3: Participation in post-compulsory education and training	Recent training	ind.05
			Lifelong learning (employees)	ind.06
			Lifelong learning (aged 25-64)	ind.07
			VET students	ind.08
			Training deficit	ind.09
		IG4: Attainment from post-compulsory education and training	Tertiary attainment	ind.10
			High computer skills	ind.11
2 P2: Skills Activation	SP3: Transition from education to work	*(IG missing)*	Early leavers from training	ind.12
			NEETs	ind.13
			Recent graduates in employment	ind.14
	SP4: Activity rates		Activity rate (aged 15-24)	ind.15
			Activity rate (aged 25-54)	ind.16
			Activity rate (aged 55-64)	ind.17
3 P3: Skills Matching	SP5: Unemploy-ment and vacancies	IG5: Unemployment	Long-term unemployment	ind.18
		IG6:Vacancies	Structural vacancies	ind.19
		IG7:Under-employment	Underemployed part-time workers	ind.20
	SP6: Skills mismatch	*(IG missing)*	Skills obsolescence	ind.21
			Higher education mismatch	ind.22

European Skills Index version 2018

Indicator (ind)		Sub-pillar (SP)	Pillar (P)
Pre-primary pupil-to-teacher ratio	ind.1	SP1: Compulsary education	1 P1: Skills Development
Upper secondary attainment (aged 15-64)	ind.2		
Reading, maths & science scores (aged 15)	ind.3		
Recent training	ind.4	SP2: Post-compulsary education and training	
VET students	ind.5		
High computer skills	ind.6		
Early leavers from training	ind.7	SP3: Transition from education to work	2 P2: Skills Activation
Recent graduates in employment	ind.8		
Activity rate (aged 25-54)	ind.9	SP4: Activity rates	
Activity rate (aged 20-24)	ind.10		
Long-term unemployment	ind.11	SP5: Unemploy-ment and vacancies	3 P3: Skills Matching
Underemployed part-time workers	ind.12		
Higher education mismatch	ind.13	SP6: Skills mismatch	
ISCED 5-8 proportion of low wage earners	ind.14		
Qualification mismatch	ind.15		

Notes: Making Skills Work Index (left) was an earlier beta-version of the European Skills Index (right). Eleven indicators (in red, left table) were either removed or replaced with four indicators (in green, right table).

Source: European Commission, Joint Research Centre, 2018 (based on the European Skills Index report).

Populating the framework with indicators

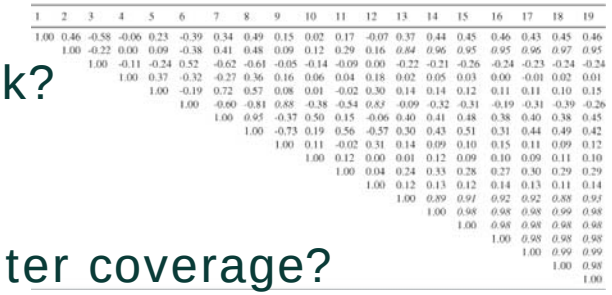
An iterative process!

⇔ Conceptual framework

⇔ Statistical properties of the indicators

➔ ***See sessions on Statistical Coherence***

- does the **correlation structure** reflect the conceptual framework?
  - If not, would changing the specification of the indicator (i.e., denominator) make a difference?
  - Is data coverage acceptable? Is there another proxy with better coverage?
  - Are the latent dimension(s) confirming the conceptual structure?



⇔ How can we interpret the correlation or principal component analysis (PCA) outcomes in light of the conceptual framework?

- Often a reason to refine/rethink the indicator framework (consider indicator development as a **learning process**)

Populating the framework with indicators

…an iterative process! ⟲

⇔ Does it meet the expectation of experts, analysts, policy users?

- Stakeholders' acceptance is important

➔ Participatory development process

- Helps articulate and refine different perspectives
- Compromises & normative choices unavoidable
  => these should be well documented!

The Developer's headache...

- Data (source, type, denominators, etc.)
- IT platforms
- Structure…
- Tradeoffs…

What IT platform to use?

- Excel: ***→ see session on COIN Tool***
  - "WYSIWYG": offers quick assessment of data quality profile; likely to share results (graphs); available everywhere
  - easy to lose track of manipulations – *make sure to use functions & keep dynamic links to original data sources*

- Statistical software, i.e. STATA, Matlab or R (or Excel VBA)
  - Less intuitive, high initial learning cost –
  - Easy to document choices in script languages (i.e., stata .do files)
  - Excel not ideal for some steps (PCA, simulations, etc…)
- Structuring data: *find layout most suitable for context*
- Downloads can be programmed (see **`readSDMX`** for R; **`getdata`** of STATA; etc.); bulk download preferable also in Excel
- *Keep track of numerators, denominators, different versions tested*

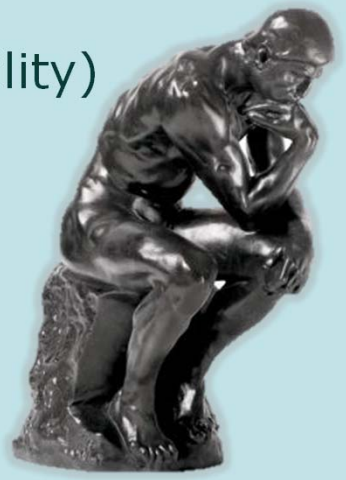
What data to use?

A. Use available data
- Was it collected for a similar purpose?
- Was it collected for another purpose but is relevant?

B. Generate new data
- Run surveys => costs; coverage; replicability
- Build from microdata => cost (also of replicability)
- Exploit Big Data (or admin data)
  => If desired indicators or desired granularity not available (cost, replicability)

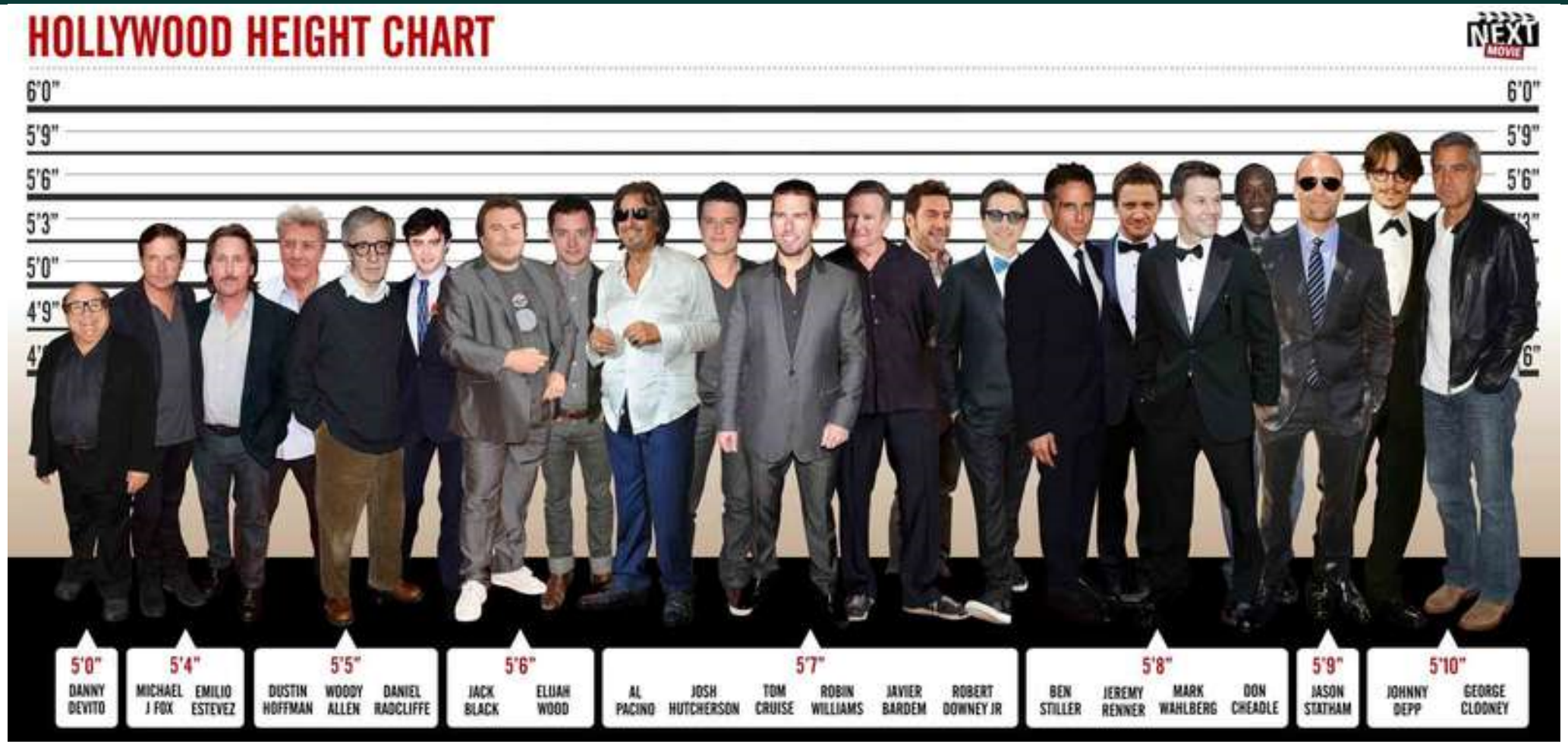
C. Combine different sources
- Consider costs (€, time), ease of interpretation (intuitive?)
- Spell out the desired quality for the indicators to collect!
  - Can distinguish country performance? Missing data acceptable?

Avoid GI-GO



Select meaningful indicators

Fit for purpose? Can it distinguish performance?



European Commission

Choices and trade-offs

- timeliness ⇔ completeness
- quality ⇔ breadth of coverage
- Novelty ⇔ acceptance
- Sophistication ⇔ intuition

	Years											
Indicator	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
HICIT	x	x	x	x	x	x	x	x	x			
TOPINST						?	?	?	?	?		
PCT	(x)	(x)	(x)	(x)	(x)	(x)	x	x	x	x	x	
ERC									(x)	x	x	x

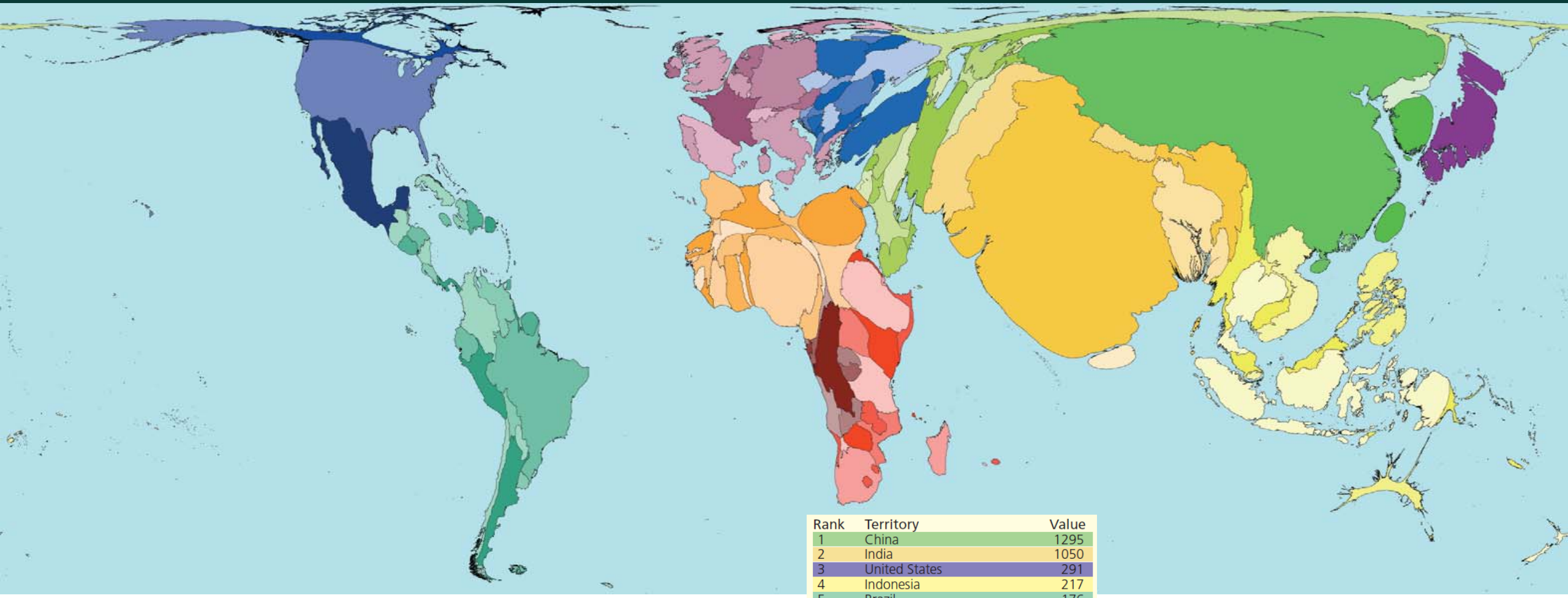
	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
HICIT	x	x	x	x	x	x	x	x	x			
TOPINST						?	?	?	?	?		
PCT	(x)	(x)	(x)	(x)	(x)	(x)	x	x	x	x		

- Consider…
  - reproducibility
  - meaning of indicators at different levels of aggregation

- How to compare performance of countries of different size?
  - Choice & interpretation of a denominator is not straight forward!

European Commission

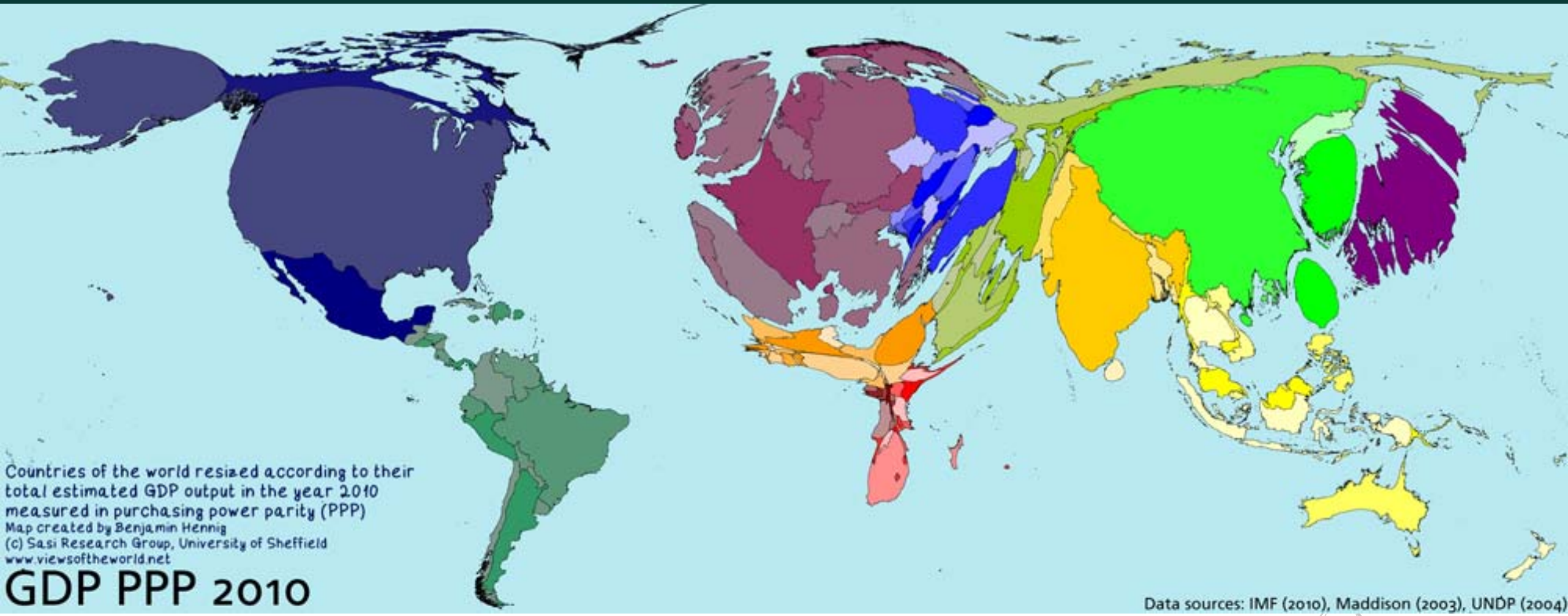
Normalize by Population...?



Rank	Territory	Value
1	China	1295
2	India	1050
3	United States	291
4	Indonesia	217
5	Brazil	176
6	Pakistan	150
7	Russian Federation	144
8	Bangladesh	144
9	Japan	128
10	Nigeria	121

millions

...or size of the economy (e.g. GDP)?



European Commission

Updating composite indicators

- Every new edition is an opportunity to refine framework & indicators
  - “Agora model”: changes and refinements help better understand the phenomenon of interest; discovery of new aspects

- Tradeoff between continuity & refinement
  [advocacy & analytical functions]

- For the developer: think of future updates at the start
  - Also in terms of data management

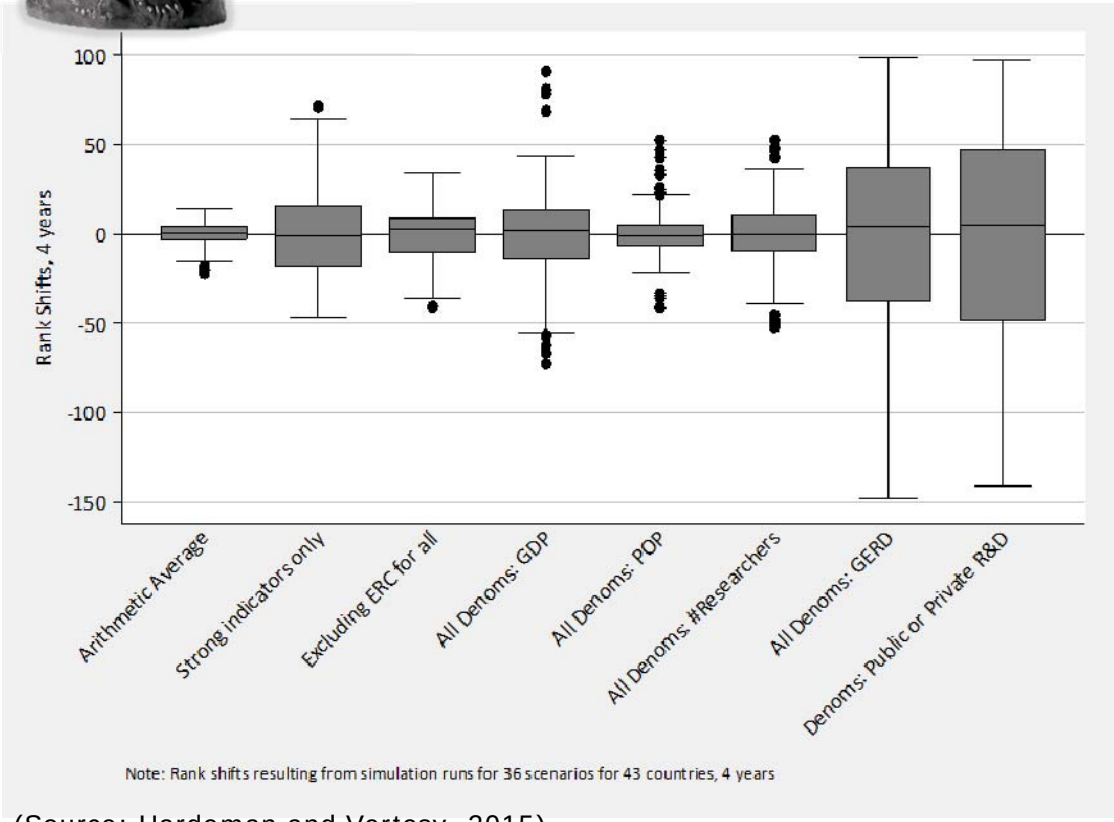
Selecting indicators

- ✓ Check the quality of available indicators
- ✓ Discuss strength and weakness of selected indicators
- ✓ Provide a summary table of key characteristics
  - Coverage (across time & space)
  - Source
  - Type (hard or soft measures; input / process / output?)
- ✓ Make your choices clear for ALL (including yourself!)
  - What, Why, How…? & What ***not***, why ***not***, how ***not***…?
- ✓ Make your indicator time-resistant (socio-political context may change!)
- ✓ Clear documentation is essential

Are you certain you have the right measure?

The impact of conceptual and methodological choices **can be quantified**
➔ ***see uncertainty & sensitivity analyses***

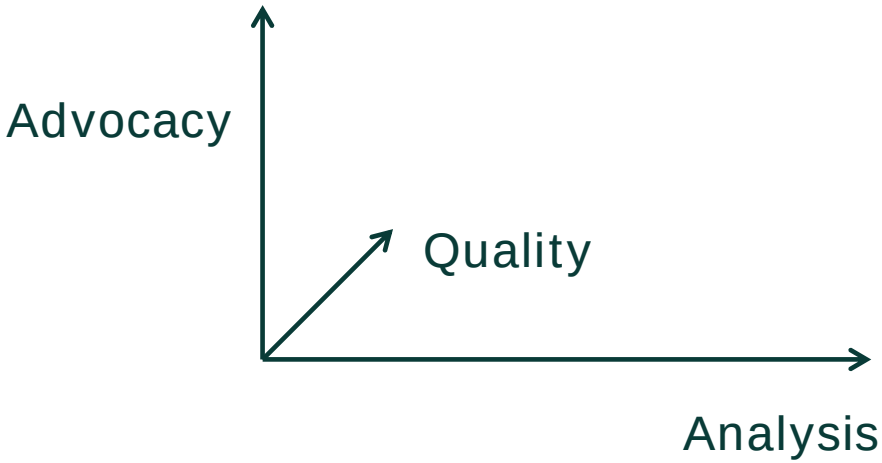
- How important are underlying assumptions?
- Identify key modelling choices
- Test their impact on the final composite scores and rankings



(Source: Hardeman and Vertesy, 2015)

Assessing
The quality of indicators

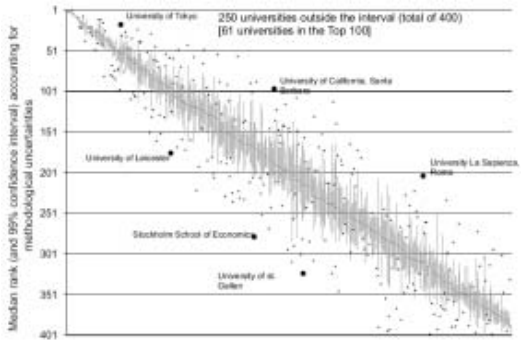
The Quality of Composite Indicators



See: Saltelli (2007)

Codified and continuously refined methodology

- The OECD-JRC Handbook (JRC-OECD, 2008)
- Audits – robustness and sensitivity analyses (i.e. Saisana et al, 2011; Paruolo et al, 2013)



European Commission

Quality profile of [composite] indicators

- Quality = accuracy?
  [closeness of computation to the “exact true values”]
- Quality = fitness for use?
  [depends on user needs, values, priorities]

- **Quality assessment frameworks:**

for Official Statistics

for Composite Indicator Frameworks

European Commission

1. The UNSD's National Quality Assurance Framework

Statistics Division

Towards a National Quality Assurance Framework: the UN Statistical Commission Initiative

- Refers to individual indicators, but relevant also for their combinations…
- Focus both on development process & results

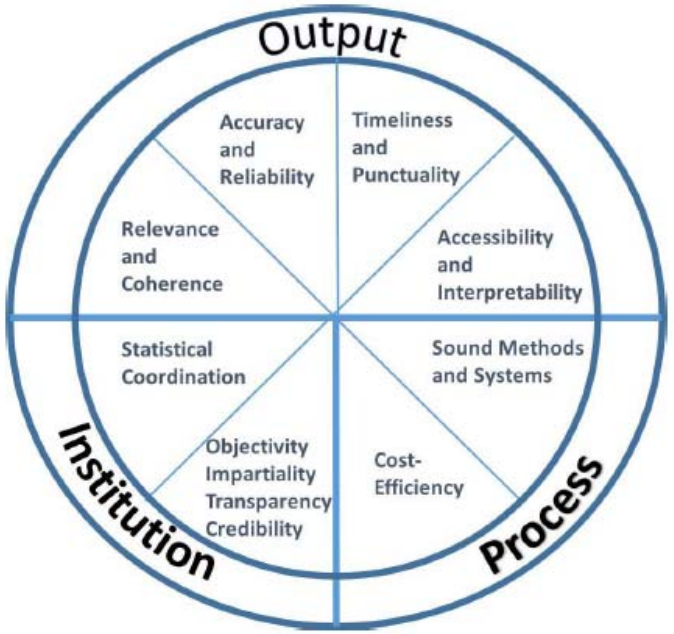


Figure 1. Quality Dimensions

Managing statistical processes

- Assuring methodological soundness
- Assuring cost-effectiveness
- Assuring soundness of implementation
- Managing the respondent burden

Managing statistical outputs

- Assuring relevance
- Assuring accuracy and reliability
- Assuring timeliness and punctuality
- Assuring accessibility and clarity
- Assuring coherence and comparability
- Managing metadata

European Commission

2. European Statistics Code of Practice

Code of Practice (CoP) & ESS Quality Assurance Framework (QAF)

15 principles focusing on the

- Institutional Environment,
- Statistical Processes,
- **Statistical Output [11-15]**

11. **Relevance:** meet the needs of users
12. **Accuracy and reliability:** accurately and reliably portray reality.
13. **Timeliness and punctuality:** released in a timely and punctual manner
14. **Coherence and Comparability:** consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources.
15. **Accessibility and Clarity:** presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

See: http://ec.europa.eu/eurostat/web/quality/european-statistics-code-of-practice

European Commission

3. The “Gisselquist framework”

Blog
What Does Good Governance Mean?

Rachel M. Gisselquist

Rachel M. Gisselquist

A framework of 10 questions to guide the development and evaluation of composite indexes
(Developed for governance indicators)

Focuses on:

- Concept, definition, operationalization
- + Data + Quality of methodology

Critical Questions:

1. What precisely does it aim to measure?
2. Content validity: does the operational definition capture the concept?
3. How good (reliable, valid and complete) are the data used?
4. Is the measure (including all of its sub-components) transparent and replicable?
5. How sensitive and robust is the measure to different data and design choices?
6. Does the measure allow the analyst to address key questions of interest?

Less Critical Questions:

7. Does the measure fully capture [the concept of interest] in all its complexity? [descriptive compl.]
8. Does the measure behave as theory predicts? [theoretical fit]
9. How precise are index values and are confidence intervals specified? [precision of estimates]
10. Is the weighting ‘correct’?

(Gisselquist, 2014)

European Commission

Quality of Composite Indicators

- Fitness for purpose
- Statistical & conceptual – technical and normative aspects hard to separate
- Assessment frameworks help analyze developers' choices made with respect to:
  - the concept;
  - the operationalization process;
  - accounting for information loss

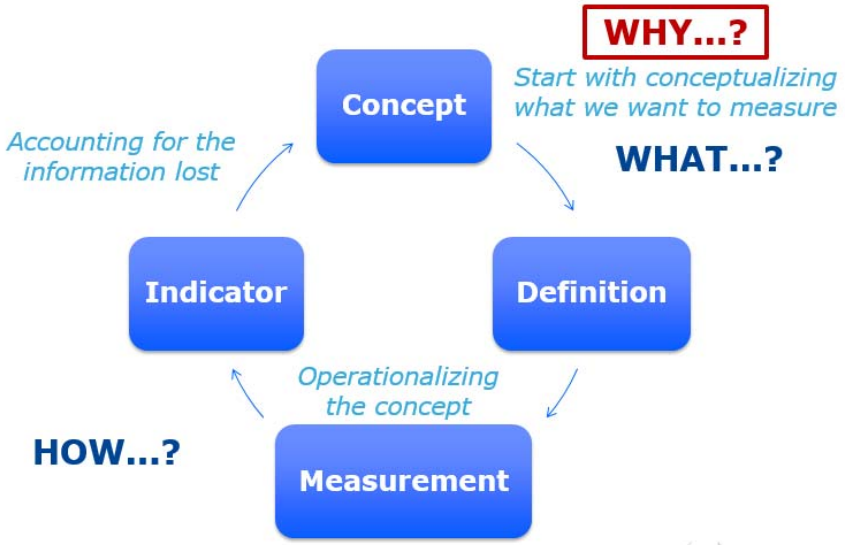
  ...and thus help interpret results

=> audit of composite indicators and frameworks

Indicators for Policy: The normative aspect

- Quantification (modelling) involves making **normative choices** about…

- Normative choices affect:
  - the concept;
  - the operationalization;
  - accounting for information loss



⇒ Composite and stand-alone indicators are alike

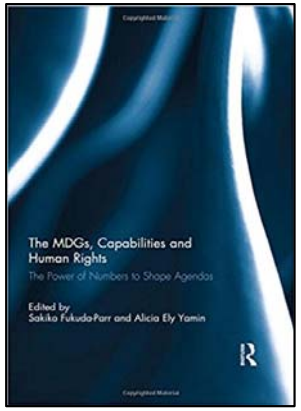
Goodhart's Law

When a measure becomes the target, it ceases to be a good measure...

Unleashing the Power of Numbers

European Commission

The power of numbers



Case of the Millennium Development Goals (MDGs)

*"While quantification is the key strength of global goals, it also involves **simplification**, **reification** and **abstraction**, which have far-reaching implications for redefining priorities"*

*Setting MDG goals/targets influenced policy priorities and had **normative effects** on development discourses;*

*All MDG goals/targets "led to **unintended consequences** in diverting attention from other important objectives and reshaping development thinking"*

(Fukuda-Parr, S., Yamin, A.E., Greenstein, J., 2014)

European Commission

Indicators are shaped by policy needs & discourse
Indicators, in turn, influence policy discourse

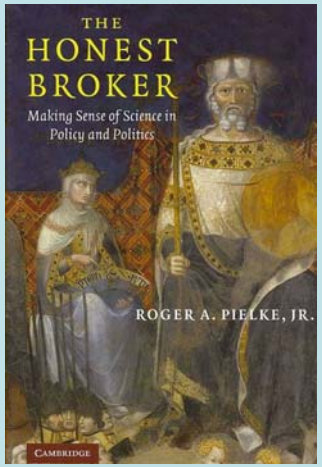
Indicators are embedded in a socio-political context

- Indicators & indicator frameworks are **value laden**; reflect policy discourse (Godin, 2002)

- Indicator developers & users should be aware of the consequences:

  - The “agora model” (Barré, 2001, 2010): indicators are debating devices – it’s the process that matters!

  - Be an “Honest broker” (Pielke, 2007)



European Commission

Be an honest broker with Composite Indicators

Use available tools to **increase robustness and credibility**:

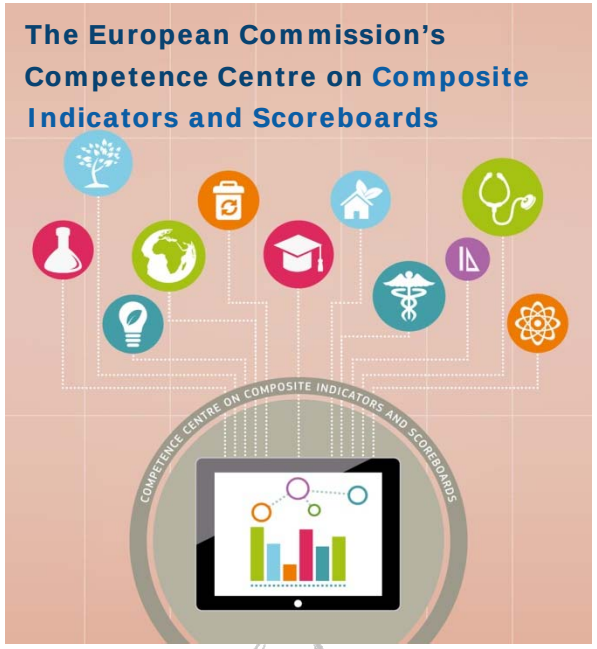
1. Transparency — detailed description of methodology, data sources, assumptions
2. Statistical soundness — analysis of correlations, data structure, effects of weights, etc.
3. Uncertainty and sensitivity analysis — check effect of alternative but plausible assumptions. Honestly acknowledge uncertainty.

THANK YOU

Welcome to email us at: daniel.vertesy@ec.europa.eu
jrc-coin@ec.europa.eu

COIN in the EU Science Hub
https://ec.europa.eu/jrc/en/coin

COIN tools are available at:
https://composite-indicators.jrc.ec.europa.eu/



European Commission

References / 1

Barré, R. (2001). The Agora model of innovation systems: S&T indicators for a democratic knowledge society. *Research Evaluation*, *10*(1), 13-18.

Barré, R. (2010). Towards socially robust S&T indicators: Indicators as debatable devices, enabling collective learning. *Research Evaluation*, *19*(3), 227-231.

Cedefop (2018) European Skills Index Technical Report. URL: https://skillspanorama.cedefop.europa.eu/sites/default/files/ESI%20-%20Technical%20Report.pdf

Davis, K. E., Fisher, A., Kingsbury, B. and Merry, S.E. (eds.) (2012) *Governance by Indicators: Global Power Through Quantification and Rankings.* Oxford: Oxford University Press.

Ferretti, F., Guimaraes Pereira, A., Hardeman, S. & Vertesy, D. (2018) "Research Excellence Indicators: time to re-imagine the 'making of'?" *Science and Public Policy 45*(5) 731-741

Fukuda-Parr, S., Yamin, A. E., & Greenstein, J. (2014). The power of numbers: a critical review of millennium development goal targets for human development and human rights. *Journal of Human Development and Capabilities*, *15*(2-3), 105-117.

Gisselquist, R. (2014) Developing and evaluating governance indexes: 10 questions. *Policy Studies 35*(5),513-531

Hardeman S., Van Roy V., Vértesy D. and Saisana M. (2013) *An analysis of national research systems (I) A Composite Indicator for Scientific and Technological Research Excellence* EUR 26093. Luxembourg: PO EU; doi: 10.2788/95887

Hardeman S. & Vértesy D. (2015) *An update of the research excellence indicator.* Luxembourg: Publications Office of the EU; doi: 10.2760/770103

References / 2

JRC-OECD (2008). *Handbook on constructing composite indicators: methodology and user guide*. Paris: OECD.

Paruolo, P., Saisana M. and Saltelli A. (2013) “Ratings and rankings: voodoo or science?” *Journal of the Royal Statistical Society: Series A*. 176(3), 609–634

Power, M. (1994). *The audit explosion* (No. 7). Demos Paper No. 7.

Rottenburg, R., Merry, S.E., Park, S-J. and Mugler, J. (eds.) (2015) *The World of Indicators: The Making of Governmental Knowledge through Quantification.* Cambridge: Cambridge University Press

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research*, 81(1), 65-77.

Saisana M., D’Hombres B., Saltelli A. (2011) “Rickety Numbers: Volatility of university rankings and policy implications.” *Research Policy* 40, 165–177

Saisana M., Saltelli A., Tarantola S. (2005) “Uncertainty and sensitivity analysis techniques as tools for the analysis and validation of composite indicators”, *Journal of the Royal Statistical Society A* 168(2), 307-323.

United Nations (2012) *National Quality Assurance Frameworks*
[URL: http://unstats.un.org/unsd/dnss/qualityNQAF/nqaf.aspx]