

# You say you want a [data] revolution: Reflections one year on

Bojan Nastav<sup>a</sup> and Steve MacFeely<sup>b,c,\*</sup>

<sup>a</sup>Statistical Consultant, Ljubljana, Slovenia

<sup>b</sup>United Nations Conference on Trade and Development, Geneva, Switzerland

<sup>c</sup>Centre for Policy Studies, University College Cork, Cork, Ireland

**Abstract.** In August 2019, the IAOS discussion platform was launched with a special session at the ISI World Congress in Kuala Lumpur, Malaysia. The first paper published as part of that discussion series was ‘*You say you want a [data] Revolution: A proposal to use unofficial statistics for the SDG Global Indicator Framework*’. This follow-on paper looks back over the year since the launch, summarizes the online debate, highlights some other relevant papers, and reflects on where the discussion rests today.

Keywords: 2030 Agenda, accreditation, risk management, UN World Data Forum, data revolution

## 1. A lot can happen in a year

In August 2019, the IAOS discussion platform was launched with a special session at the ISI World Congress in Kuala Lumpur, Malaysia. It was our great privilege to have had our paper ‘*You say you want a [data] Revolution: A proposal to use unofficial statistics for the SDG Global Indicator Framework*’ [1] selected for the inaugural discussion. We would like to thank the Statistical Journal of the IAOS for this great honour.

The discussion platform was launched, and the paper published in the halcyon days before COVID-19. The pandemic triggered an escalation in the already growing demand for new and real time statistics, serving to highlight further the challenges facing official statistics. The pandemic also saw a noticeable step-up in statistical information being provided by unofficial sources, not least, Google’s COVID-19 *Community Mobility Reports* [2] and the Deep Knowledge Group [3] with their AI generated ‘*COVID-19 Regional Safety Assessment*’.

A year has passed since the original paper was published. This short paper gives a brief synopsis of the

debate that took place on the discussion platform and our reactions or response. Before that, the paper begins by quickly outlining some issues we forgot, and some things we quickly realized.

## 2. Things we forgot

No sooner was the digital ink dry on the original paper when we realized there were at least three important omissions. Firstly, in the context of the Sustainable Development Goal indicators, we had not mentioned or addressed the possible outcomes of the 2020 SDG Comprehensive Review. This review was the first of two, envisaged during the 2030 Agenda process, whereby, non-performing Tier 2 and 3 indicators could be culled or replaced by superior indicators. The first comprehensive review was scheduled for 2020 the second for 2025. The 2020 review, signed off by the UN Statistical Commission in March 2020 led to quite a number of significant reclassifications (see Table 1). When the original paper was published (using IAEG-SDG updates from April 2019), only 44% of indicators were Tier 1 and 39% were Tier 2. By July 2020, an additional 22 indicators had been reclassified as Tier 1. Furthermore, Tier 3 indicators had been eliminated completely. Table 1 shows the progress made by the global statisti-

---

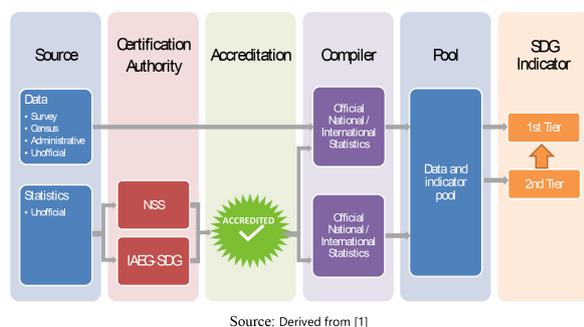
\*Corresponding author: Steve MacFeely, United Nations Conference on Trade and Development, Geneva, Switzerland. E-mail: steve.macfeely@un.org.

Table 1  
Number of SDG indicators by Tier (updated to July 2020) Source: Derived from IAEG-SDG Tier Classification [4]

Tier classification	December 2016		December 2017		December 2018		December 2019		July 2020	
	Number	%	Number	%	Number	%	Number	%	Number	%
1	81	35	93	40	101	44	116	50	123	53
2	57	25	66	28	84	36	92	40	106	46
3	88	38	68	29	41	18	20	9	–	–
Multiple	4	2	5	2	6	3	4	2	2	1
Total	230	100	232	100	232	100	232	100	231	100

cal community, year by year, in particular, developing new statistical concepts and methodologies.

The elimination of Tier 3 indicators of course means that our original schematic can be simplified somewhat – see Fig. 1.



Source: Derived from [1]

Fig. 1. Simplified proposed future: Using unofficial data and statistics to compile SDG indicators.

Another lacunae in the argument presented in our original paper is what would happen if a country objected to a globally accredited indicator or to particular estimated 'value' for their country? Obviously, this wouldn't be an issue at national level, but it is a very real possibility at global level. Although we had, to some extent, highlighted this risk in 'Who measures' [1, Section 2.2] where we had noted that countries, understandably protective of their reputations, can be sensitive about what is measured and who does the measurement. But we had not developed any argument or mechanism as to how this should be addressed. The purpose of this paper is not to extend the original paper further, but it seems that some sort of disputes reconciliation mechanism would be required.

Finally, another issue, indirectly referenced but not explicitly dealt with, was citizen science. This particular 'movement' for want of a better word, has been rapidly gaining traction and is likely to become an important player in national and global data ecosystems. But incorporating citizen science may pose some problems for the proposal as we originally outlined it. In the original paper, we had argued, as a basic condition, that data

must be available for the period 2015–2030. However, in such a rapidly developing field, this may prove too restrictive for some citizen science data. So, a lingering question, assuming our original proposal found favour, is whether this condition could be relaxed?

### 3. Things we subsequently realised

Quite soon after the paper was published, we realized that many readers had not understood the significance of the title and had not made the connection with the famous 1968 Beatles song 'Revolution' that reflected their mixed feelings on that turbulent year. Anyone who did make the connection and understood the reference to the opening lines of the song 'You say you want a revolution/well, you know/we all want to change the world' hopefully also understood that the ambivalence expressed by the Beatles when they sang 'don't you know that you can count me out/in' well reflected our own uncertainties regarding events in the data world and doubts as to how we as official statisticians should respond. As we noted in the introduction of the original paper, although we were convinced that no additional funding would be diverted towards national and international statistical systems, and that other solutions were needed, we nevertheless put forward our proposal somewhat reluctantly. In other words, like the Beatles, you could count us in/out.

### 4. The online discussion

As noted in the introduction, the discussion began at the launch of the discussion platform at the 2019 ISI World Congress. At this special session, we had the opportunity to present our argument, and then debate the pros and cons with a lively and experienced audience. Predictably enough, the argument divided the room and proved to be quite a good barometer of the debate that would follow on the platform itself.

To prompt discussion, the online discussion began with four 'provocative statements' (see appendix). The

online debate, such as it was, began in late August and continued through September 2019, then went dormant until February 2020, and then ended after a brief resumption in June. The online discussion can be categorized broadly, into four cohorts: those opposed; those in support; those who had questions; those who argued against proposals we didn't actually make – we have termed these misunderstandings. Each of these are summarized below:

**Opposed:** naturally, there were those who didn't like the proposal. The concerns they raised were legitimate, and ones we had largely anticipated and grappled with in 'Risks associated with adopting this proposal' [1, Section 3.3] of the original paper. The first concern or objection was that our proposal runs the risk of bypassing national statistical offices, and thus weakening rather than strengthening national statistical systems. Allied to this, some didn't like the franchising model as they felt it would make the existing situation worse. There were also several concerns that a certification process would be too heavy, and resource intensive, and might distract from the core business of producing statistics. Others raised the particular challenges faced by developing countries, and whether this was really feasible for NSOs that are badly under resourced to begin with.

**Support:** others were more positive towards or supportive of the proposal. In general, their support stemmed from a recognition that the Internet of Things, the data revolution and citizen science will all, most likely, make a range of new statistics available, so why not use them? Others argued that NSOs already vet and use a wide range of unofficial data, so widening the scope to include statistics and formalizing this process might not be a huge step. Others gave guarded support, but in doing so warned that accreditation should be used to legitimize the use of unofficial statistics, and not applied to prevent their use i.e. the process should not be used to create 'non-tariff barriers' against unofficial statistics.

**Questions:** Some very good questions were posed in the debate. The first question asked whether this was simply an academic exercise, or had we identified examples of unofficial statistics or indicators that could be used to populate the SDG Global Indicator Framework? Another question, perhaps the most important one asked, and one we had agonized over ourselves, was how to assess the motive of the compiler; how can a NSO or any homologation authority assess whether an unofficial compiler has 'skin in the game' i.e. if they have a vested interest or not? There was a question mark

too around the incentives for producers; why would a compiler of an unofficial statistic subject themselves to a certification process (and the embarrassment of possible rejection)? Finally, others wondered how heavy or complex would the vetting or accreditation process need to be.

**Misunderstandings:** unsurprisingly, and despite our best efforts, our paper wasn't always as clear as it should have been. As a consequence, it seems some misunderstandings arose and some commentators responded to proposals we didn't make (or rather hadn't intended to make). One principal misunderstanding centred on the distinction between data and statistics. Our paper proposed the accreditation of unofficial statistics, not data, which we recognise are already used extensively by NSOs and international organisations. There also appeared to be some misunderstandings regarding the role of the SDG Global Indicator Framework and the distinction between national and international statistics more generally.

## 5. Parallel discussions

In an attempt to encourage participation on the online discussion, we made several presentations to different groups. The chair of the Advisory Committee to the 'Accountability for gender equality in education' at the University College London Institute of Education generously allowed us to present our arguments in January 2020. That same month we also were invited to present our paper to the Citizen Science 'WeObserve SDGs Community of Practice'. At the margins of the United Nations Statistics Commission in March 2020, we also presented our case to a large Civil Society group. During the year, several NSOs or representatives of a number of NSSs contacted us to discuss the paper. Many seemed interested, dare we say enthusiastic, but unfortunately very few followed up on the promises to post their reactions, positive or negative, on the online discussion platform. See Section 7 below.

A separate online debate began on LinkedIn in the lead up to the World Data Forum in October 2020, in reaction to a blog that summarized our paper posted on UN Brief.<sup>1</sup> One important comment felt our approach was too binary and argued for a more traffic light type approach. In the original paper we had raised this issue

---

<sup>1</sup>[https://theunbrief.substack.com/p/a-data-revolution-for-sdg-indicators-397?r=4ehvo&utm\\_campaign=post&utm\\_medium=web&utm\\_source=linkedin](https://theunbrief.substack.com/p/a-data-revolution-for-sdg-indicators-397?r=4ehvo&utm_campaign=post&utm_medium=web&utm_source=linkedin).

ourselves, noting that outside the narrow SDG purpose, a less binary approach that includes intermediate or experimental certification would probably make sense.

The paper was also presented at two sessions at the 2020 World Data Forum, in sessions combined with the Citizen Science (TA2.01 – Integrating Citizen Science into the Official SDG Monitoring Mechanisms and A Proposal to Use Unofficial Statistics for SDG Reporting (to Deal with Crises) and Geo-Spatial Communities (TA5.03 – Learning from geo open data platforms that worked in the midst of a pandemic). An interesting feature of the geo-spatial discussion was they had taken our proposal and refined it further. It will be interesting to see whether this idea is developed further.

## 6. Our reaction

Firstly, we would like to thank all of the commentators for taking the time to read and react to our paper. From the face-to-face meetings and online discussions it is very difficult to assess whether the majority are for or against the proposal. The numbers engaged have not been sufficiently large to make a determination. Perhaps too (as there was no vote per se) those who agreed with the statements felt no need to jump in and comment. The concerns expressed, came as no surprise; on the contrary, we had articulated most of the same fears and concerns ourselves in the original paper. Still, it was somehow reassuring to know we hadn't missed anything really obvious. Furthermore, we empathise with those concerns. So, rather than attempt to rebut them, we simply re-ask, what should we do then? It's a valid question to ask, as a notable feature of the debate, was that no one challenged or remarked on our assessment of the situation, only our proposed solution. Discussing official statistics, Rolland [5, p. 373] argued that as competitors move in to fill the gaps between demand and supply, official statistics is faced with the threat of obsolescence, and thus concluded that 'the challenges make it urgent to rethink the foundations of the industry'. We fear Rolland may be correct. This was our attempt to rethink the foundations and perhaps turn competitors into collaborators

One very important issue raised, is the statistical capacity of developing countries. In retrospect, we realise that our argument was perhaps too 'developed country' centric, and that we did not give sufficient attention to the plight of NSOs in developing countries or consider their particular circumstances. That said, this would not invalidate the global argument, or for coun-

tries that have the resources or capacity to implement such a system. But it's an important point nevertheless.

As noted above, some excellent questions were put to us during the discussion. The question on whether there are examples of unofficial statistics or indicators that could be used to populate the SDG Global Indicator Framework, is best answered by Fraisl et al. [6], who have systematically identified and mapped examples across all 17 SDG goals, indicator tiers and indicator custodian agency where citizen science can be used to populate SDG indicators.<sup>2</sup> Their assessment, summarised in Fig. 1, suggests there is a great number of opportunities to utilise citizen science data. We are not aware of any comparable, comprehensive studies for civil society, academic, NGO or private sector data, but from past experience we expect we would find similar, if less extensive, opportunities. For example, in 2016 UNCTAD [7] was the first UN agency to publish a statistical analysis of the SDGs. That study used a wide variety of unofficial statistics to provide a benchmark assessment of the 2030 Agenda. The citizen science argument is gaining traction, with mainstream media asking if for 68% of environmental indicators there is not enough data to assess progress [8] then why are we not trying new approaches [9]?

A more difficult question to answer is whether an accreditation mechanism can ensure that the purpose of any statistic is to provide impartial information and is not designed to support or advocate for a particular ideology or cause. While we anticipated this problem, noting that 'conflict of interest is always a risk when consumers of data become compilers' [1, p. 319] we did not put forward a specific solution. But clearly, this is something that needs to be considered carefully. In the original paper we discussed the challenges surrounding technocratic objectivity, as official statistics already play the dual role of both measuring and shaping reality, posing some knotty philosophical questions as to what neutrality really means [10,11]. But the point is well made, and although we are aware of it, we don't have any easy solution to propose.

On the incentive question, perhaps optimistically, we assumed that NSOs in a national context, and the United Nations Statistical Commission in the global

---

<sup>2</sup>The authors also note that integrating citizen science into official statistics may not be straight-forward. They suggest a number of steps that might help, including, awareness building, developing case studies, setting out criteria for ensuring data quality and procedures, integrating methodologies, aligning definitions, and using standard classifications and other agreed formats and standards.



gument, but they are noteworthy, as these types of misunderstandings persist and are at the centre of a number of tensions between international organisations and countries [12,14].

## 7. Barriers to debate

Over the course of the 12 months that the first discussion was online, the debate was, although focused and to the point, less active than we had hoped for. Obviously as authors we must take some responsibility for this; evidently our paper wasn't as urgent or as engaging as we had hoped. But from offline discussions, it is clear that other barriers to debate exist too. The first is nothing new or surprising; some are simply reluctant to engage to public debate as they feel they are insufficiently expert or confident in their views to express them in public. Others are reluctant to engage with online debates, as they have witnessed bad online behavior or bullying and understandably have no wish to be on the receiving end such abuse. Of course, a professionally run discussion platform, such as the IAOS platform, has guidelines and standards to deal with such occurrences, and is monitored carefully by the IAOS journal editor, but it seems the concern is abroad. These concerns are understandable. Another concern, expressed privately, was more surprising. In the aftermath of the decision in Australia, where a public servant was dismissed for tweeting in breach of the Australian Public Service (APS) Code of Conduct [15], some expressed concern that this aggressive stance could spread to other countries, and perhaps be applied retrospectively. We feel this is a worrying development. Every official statistician understands the need for impartiality and political neutrality, and the need to respect their contractual obligations, but this surely should not embargo debate on issues that are of central importance to our profession.

## 8. Conclusion

We would like to conclude this article, in much the same way we concluded the original paper. The Dubai Declaration, drafted at the conclusion of the 2018 UN World Data Forum acknowledges 'that the data demands for the 2030 Agenda require urgent new solutions that leverage the power of new data sources and technologies through partnerships between national statistical authorities and the private sector, civil society, and the academia and other research institutions' [16,

p. 7]. The UN Statistics Division, summarizing the debate of a special session at the 50th session of the UN Statistics Commission in 2019 on 'The Future of Economic Statistics', state 'In an apparent paradigm shift, there is now an overriding consensus that both developed and developing statistical system should actively pursue and accelerate the integration of these new data sources in their statistical products and services [17, p. 2].' They also noted 'Statistical agencies alone do not have the capacity to meet these challenges, they need to work with academia, private sector agencies, NGOs and other government agencies to obtain the data, tools and know-how required [17, p. 3].' So there appears to be general acknowledgement that greater collaboration is the way forward.

But as Carranza [18, p. 19] notes 'Although civil society organizations are also strategically welcomed by the "A world that counts" declaration to collaborate in aiding NSOs in the quest of filling data gaps, many tactical and operational issues are not immediately recognized in the walk to reality.' This applies, not only to civil society statistics, but arguably to all unofficial statistics, whether civil society, citizen science, non-governmental, academic or private sector. Greater collaboration is the way forward – we just haven't figured out how to operationalize it yet. Perhaps by not putting an accreditation system in place we have created barriers to entry rather than the other way around. Perhaps we, official statisticians, are more protectionist than we care to admit? But maybe the winds of change are starting to blow, gently. A recent article from Ricciato et al. [19] entitled '*Trusted Smart Statistics: How new data will change official Statistics*' discusses the need for a more inter-disciplinary participatory approach to official statistics, one that includes contributions beyond the 'traditional competence perimeter of statisticians' towards what they call 'Citizen Statistics'.

We leave the last word to Gennari and Navarro [20], who in a no-holds-barred polemic argue we need to 'rekindle the data revolution' saying 'A more flexible yet institutional approach is needed to accommodate the needs of different actors while maintaining coherence and the involvement of official data producers. Undoubtedly, this includes the private sector. Alternative data sources are usually owned by private companies, but data exchange protocols are often missing. We need to address the mutual distrust and the complex transactions between the private sector and national statistical systems pragmatically rather than dogmatically'.

## References

- [1] MacFeely S. Nastav B. You say you want a (data) revolution? A proposal to use unofficial statistics for the SDG Global Indicator Framework. *Statistical Journal of the International Association of Official Statistics*. 2019; 35(3): 309-327.
- [2] Google (2020). Community Mobility Reports. Available at: <https://www.google.com/covid19/mobility/>; [last accessed: 22.06.2020].
- [3] Deep Knowledge Group (2020). COVID-19 Regional Safety Assessment. Available at: <http://analytics.dkv.global/covid-regional-assessment-200-regions/full-report.pdf>. [last accessed 22.06.2020].
- [4] Inter-Agency and Expert Group on Sustainable Development Goal Indicators (2020). Tier Classification for Global SDG Indicators as of April 2020. Available at: <https://unstats.un.org/sdgs/iaeg-sdgs/tier-classification/> [last accessed August 20, 2020].
- [5] Rolland A. The concept and commodity of official statistics. *Statistical Journal of the International Association of Official Statistics*. 2017; 33(2): 373-385.
- [6] Fraisl D, Campbell J, See L, When U, Wardlaw J, Gold M, Moorthy I, Arias R, Piera J, Oliver JL, Masó J, Penker M, Fritz S. (2020). Mapping citizen science contributions to the UN sustainable development goals. *Sustainability Science*. Available at: <https://doi.org/10.1007/s11625-020-00833-7> [last accessed 10.07.2020].
- [7] UNCTAD (2016). Development and Globalization: Facts and Figures 2016. Available at: <https://stats.unctad.org/Dgff2016/> [last accessed 10.07.2020].
- [8] UNEP (2019). Measuring Progress: Towards Achieving the Environmental Dimension of the SDGs. 13 March 2019. Available at: <https://www.unenvironment.org/resources/report/measuring-progress-towards-achieving-environmental-dimension-sdgs> [Last accessed 21 July 2020].
- [9] Kart J (2020). Citizen Science Provides Useful Data for Sustainable Development Goals, International Study Shows. *Forbes*. July 15, 2020. Available at: <https://www.forbes.com/sites/jeffkart/2020/07/15/citizen-science-provides-useful-data-for-sustainable-development-goals-international-study-shows/#e74362ac604e> [Last accessed: 21 July 2020].
- [10] Desrosières A. Words and Numbers: For a Sociology of the Statistical Argument. In: Saetnan AR, Lomell HM, Hammer S. (Eds.) (2010). *The Mutual Construction of Statistics and the Society*. London: Taylor & Francis Group.
- [11] Radermacher WJ. *Official Statistics 40: Verified Facts for People in the 21st Century*. Springer International Publishing, Cham, Switzerland. 2020. doi: 10.1007/978-3-030-31492-7.
- [12] MacFeely S. Measuring the Sustainable Development Goal Indicators: An Unprecedented Statistical Challenge. *Journal of Official Statistics*. 2020; 36(1): 361-378.
- [13] Everars P. Calculating the costs of products and services of the European Statistical System. Session on the Value and costs of national and international statistics, at the ISI 2017 – 61st World Statistics Congress, Marrakesh, 17–21 July 2017, organized by OECD.
- [14] Kapto S. Layers of Politics and Power Struggles in the SDG Indicators Process. *Global Policy*. 2019; 10(1): 134-136. doi: 10.1111/1758-5899.12630.
- [15] Byrne E (2019). Public servant loses free speech High Court case over tweets criticising government policies. *Australian Broadcasting Corporation News*. Updated 7 Aug 2019, 4:15 pm. Available at: <https://www.abc.net.au/news/2019-08-07/high-court-free-speech-public-service-banerji-decision/11377990> [accessed 06.01.2020].
- [16] United Nations (2018). The Dubai Declaration: Supporting the Implementation of the Cape Town Global Action Plan for Sustainable Development Data. October 24, 2018. Available at: <https://unstats.un.org/sdgs/hlg/dubai-declaration/>; [last accessed 29.06.2020]
- [17] United Nations Statistics Division (2019). Summary on the Update of the System of Economic Statistics. Friday Seminar on Emerging Issues – The Future of Economic Statistics. 1 March 2019, United Nations Headquarters, New York. Available at: [https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3f-Summary\\_Friday-Seminar-The-Future-of-Economic-Statistics-E.pdf](https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3f-Summary_Friday-Seminar-The-Future-of-Economic-Statistics-E.pdf) [last accessed 29.06.2020].
- [18] Carranza J. Citizen to government data partnerships: What can we learn from and recommend to civil society groups working in the official statistics domain? *Eurostat Working Paper Series – November 2018*. doi: 10.2785/728477.
- [19] Ricciato F, Wirthmann A, Hahn M. Trusted Smart Statistics: How new data will change official statistics. *Data & Policy*. 2020; 2: e7. doi: 10.1017/dap.2020.7.
- [20] Gennari P, Navarro DK. Are We Serious About Achieving the SDGs? A Statistician's Perspective. *IISD SDG Knowledge Hub*. 14 January, 2020. <https://sdg.iisd.org/commentary/guest-articles/are-we-serious-about-achieving-the-sdgs-a-statisticians-perspective/> [accessed 17.02.2020].

## Appendix 1 – Provocative Statements

### Main statement for discussion:

Official Statistics should consider switching from a purely production or manufacturing based model to a mixed business model: one combining the manufacture of official statistics with the franchising of production under license.

### Other subsidiary statements to be discussed are:

On the role of National Statistical Organisations (NSO's) and International Statistical Organizations (IO's)

In this post-truth era, NSOs and IOs should assert their mandate and legitimate role as custodians of knowledge and protectors of deliberative public spaces. The statistics community is underestimating the changes underway in the world of data and statistics. The various networks and philanthropic partnerships established to strengthen data ecosystems and promote collaboration in recent years may be counterproductive, inadvertently undermining the UN. Therefore, Official Statistics must adapt in a way that allows it to take some control (or at least exert more influence) over a rapidly disintegrating information landscape. NSO's and IO's must act now to ensure the integrity of Official Statistics in a rapidly disintegrating information landscape.

On the control and ownership of statistical information

There will be ideological resistance in some countries to governments collecting more data – the fear of a Big Brother state. The neo-liberal agenda aims to minimize the role of the public sector – even in the data sphere. There will be resistance to expanded government oversight and accreditation. NSO's and IO's must act now to ensure this resistance does not erode the right to live in an informed society.

On the SDG Global Indicator Framework

It is highly improbable that based on current 'Official Statistics' by 2030 the SDG Global Indicator Framework will be fully, or even close to fully, populated. Currently less than half of the framework is populated. However, the UNSC could change into the body with the authority and competence to certify statistics as 'fit for purpose' – to review unofficial statistics and assess whether they can be certified as 'Official' for the purposes of populating the SDG Global Indicator Framework. It would be what the Académie des Sciences or the Royal Society was to the Victorian era, in terms of homologation.