



21 February 2022

Chair:	Ivo Havinga
Reporter:	Clarence Lio, Jacqueline Chan
Breakout room:	1

### Participants

- **Juan Daniel OVIEDO**
- **Ivo Havinga**
- **Eric (Floridata)**
- **Brian King**
  - **National Accounts – Statistics Ireland**
- **Geum Young Min**
  - **ONS UK**
- **Wasim**
  - **WPSA - Qatar**
- **Marian Necula**
  - **National Statistics - Romania**
- **Fatima Al-Hammadi**
- **Mahinthan Mariasingham**
  - **ADB – Economic statistics**

### Breakout room slot:

Breakout room questions

- Have you worked with third party data brokers and aggregators?
  - In what areas/projects? Structural or one-off?
  - What did the collaboration look like (financial aspects, legal basis, reciprocity, ...)?
  - What were your experiences?
  - What was in it for them?
- How do you believe we can benefit from working with such third parties?
  - In what areas would that be most helpful?
  - Do you have suggestions on concrete third parties to contact?
  - Any concerns or tips?

### Takeaways discussion

- Challenges for getting data from third party data brokers/aggregators
  - For NSO
    - Statistical legislation complex – how the NSO can buy the data

- Data linkages - Can't meet somewhere in the mid ground of sharing
    - Costs too high – not sustainable in the long-term
    - General consensus - Unacceptable for official statistics to buy or pay for data
    - Data too specific – difficult to integrate with statistical framework
  - For third-party entities
    - Dilemma whether to sell or buy data sets – conflict with business practices
      - Competitors are companies such as Amazon and Alibaba who can sell sectors as by-products
- Solutions
  - Private and public sectors can work together to strengthen their data as a public good
    - Data exchange, Outsource data analysis and workflows
  - Medium term - Need to create a common agenda/framework to help facilitate the collaboration
  - Third party
    - If you can provide a template, can help bring together other global entities in the sector

#### Detailed Notes

- Poll - Why should data brokers and data aggregators collaborate with official statistics and provide data on GVC's?
  - We bring relevant knowledge and insights (37%)
  - They can supplement their own data with our data sources (29%)
  - We can help to improve the quality of their data (26%)
- Brian
  - Not much experience working with data aggregators
    - Presenters were great role models for this idea
  - Would be great for Statistical community to collaborate with such entities
- Juan
  - During the pandemic, very specific collaborations with domestic and regional logistic aggregators of information
    - Logica – Latin America – shared CPI data
      - continuous data on supermarkets on alcoholic beverages (taxes)
      - Set and integrated way to grasp the data from their databases
      - Optimized the collection of data from the organization
      - Set legal agreements
        - Challenges



- General consensus - Unacceptable for official statistics to buy or pay for data
  - Believe in the future it will be necessary
    - Need a mechanism to outsource working pipelines
    - Complexity will crush us to do all in-house
    - Latvia and Spain – paid for telecom data
      - And had assistance from the firm to analyze
- Juan Daniel
  - Find idea of Outsourcing very interesting
    - Very necessary
    - How to have a common agenda of setting the basic standards of some variables of information
    - Price quote has several dimensions regarding the logic and background of the CPI
      - Sometimes this discussion with the private sector is highly challenging
        - Produce the data in a way that cannot easily integrate with the official statistics framework
    - Long term - Need to find a way to break silos
    - Medium term - Need to create a common agenda/framework to help facilitate the collaboration
- Ivo
  - What would be the use cases
- Eric
  - Real Global chain
    - If you can provide a template, can help get other global entities
    - If you know the sector, it's easier to get everyone together

Chair:	Rob van Kan
Reporter:	Lianne Ippel
Breakout room:	2

**Participants**

- Buoy Somethea
- Ana Zegarra
- Malik Bani Hani
- Patrick Helbling
- Parvin Huseynuva
- Dan Sofio
- Julien Dupont
- Saeed Fayyez

**Breakout room slot:**

## Breakout room questions

- Have you worked with third party data brokers and aggregators?
  - o In what areas/projects? Structural or one-off?
  - o What did the collaboration look like (financial aspects, legal basis, reciprocity, ...)?
  - o What were your experiences?
  - o What was in it for them?
- How do you believe we can benefit from working with such third parties?
  - o In what areas would that be most helpful?
  - o Do you have suggestions on concrete third parties to contact?
  - o Any concerns or tips?

**Takeaways discussion**

- It is very important to seek collaborations with third parties to enrich our data.
- At the same time we have to be alert to which methodology is used.
- Better data results in better decisions and it is therefore our mutual social responsibility to engage in conversations to inform each other and spread cost
- We are not competition to each other (i.e., data brokers and NSI) as we both serve different purposes and provide different sorts of information to different parties
- Quality of our official statistics can be improved by collaborations between NSIs, international organisations and private parties
- It is important to create awareness/data knowledge
- Already in some cases NSIs already closely collaborate with private parties such as Dun and Bradstreet to support official statistics.

Chair:	Daniela Ravindra
Reporter:	Pierre Felx
Breakout room:	3

**Participants**

Sofia Rodrigues  
Ronald Jansen  
Wesley Van Den Berg  
Matjaz Jug  
Matti Paavonen  
Sanjev Solyvann  
Abeer Sheikhasan

**Takeaways**

Most of the discussion took place on Q1 : Working with 3<sup>rd</sup> party data brokers or aggregators. A list of potential Data Aggregators that we could NSOs and International organisations could work with

- **Data provider** (collects via web scraping) about companies active with a web presence. They have the heartbeat, knowing if companies are active or not to help improve Business Registers. How much companies are involved in IT or digital economy.
- **Statistics Netherlands** working with some transportation data providers and payments distributors to understand migration patterns.
- AirDND which collects data on AirBnB. Eurostat has signed long term contract with AirBnB.
- Business organisations or lobbyists that collect data from members. An example is the Canadian Real Estate Association in Canada which acquires data from local real estate boards and disseminates statistics on the real estate market across the country.
- Floridata is an example of a data aggregator that works for the floral industry but has not seen this happen in other sectors. Their goal is to help businesses who are sharing data. Data is only shared directly with the companies that share data. Universities help Floridata with models and data science.

Another interesting but inverse example is to explore selling NSO data or working with Data brokers such as trade data monitoring companies. This provides them visibility and credibility and can form a great partnership

There was a suggestion of having an international data to discuss administrative data holdings/sources and access.

A point was raised on the experience of international organisation OECD / IMF / UN and the ability to work with data aggregators. The example provided that it was difficult for international organisations to establish contracts but they could provide benefits y working with national organisations to act as an intermediary to help establish contracts.

Chair:	Barteld Braaksma
Reporter:	Gert Buiten
Breakout room:	4

#### Breakout room #1

- Several countries shared experiences with contacts with data brokers, aggregators etc. The results were mixed: sometimes this led to actual structural data deliveries, in other cases it stopped somewhere in the process of working and trying things out. In any case, these contacts require time and resources.
- Several types of firms were mentioned
  - o Companies that collect, aggregate and derive economic data in general by combining all sorts of government and other data
  - o Companies that support a specific supply chain (similar to Floridata)

- Companies that collect specific types of data, e.g. [www.dataprovider.com](http://www.dataprovider.com) that focuses on data from web scraping company websites in dozens of countries
- None of the countries present in the breakout room have legal access to this kind of data.
- In most some, a win/win approach was used, trying to make it also worthwhile for the private company without actually paying for the data. Examples of win/win situations are
  - sharing knowledge on methods that the companies can apply in their own process. An example comes from Hungary, where the NSI and a private company developed a real estate index.
  - provide insights on e.g. quality of the data of the company
  - jointly develop new methods, sometimes also in collaboration with academia. In the latter case, one of the options is that e.g. a Master student gets access to privately held data and publishes aggregated results that are useful for the NSI
  - in one case in the Netherlands, a 'barter construction' is used: the company shares data with the NSI and NSI publishes a tailor made table containing data that is useful for the company
  - a possible new development in the EU are data sharing services, where a trusted third party can act as a broker but is not the data owner.
- The feeling is that some kind of legal base is needed to secure regular delivery of this kind of data – and maybe also to speed up the negotiation process. A problem in this respect is that in many countries, the legal base for the NSI is meant for regular surveys in which companies have to deliver data on their own business and not on data on other companies that are in their databases.
- Some participants mentioned the need for methodological and quality issues when using data from sources with a non-statistical goal