Measuring the Economic Value of Data and Cross-Border Data Flows

David Nguyen\textsuperscript{1} & Marta Paczos\textsuperscript{2}

\textsuperscript{1}ESCoE & NIESR

\textsuperscript{2}ESCoE

ASSA/AEA Annual Convention
San Diego, 3 January 2020
Disclaimer

This research has been co-funded by the OECD as part of the research programme of the Economic Statistics Centre of Excellence (ESCoE).

The views expressed here do not necessarily reflect those of the ONS, OECD, NIESR or ESCoE.
The world’s most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules

The Economist, print edition, 6th May 2017
With the ongoing digitalisation of the economy the **scale and scope** of how businesses use data changes.

Not all data is created equal. It comes in many shapes and forms. We **focus on business value of data** (not public good).

Broad view on data as "unordered and unprocessed representation of any types of observations that are quantified and stored in symbols".

It’s **value depends on the information content** and not its ’volume’. The ’value of information depends on ways to monetise it’ (now or in the future).

This paper is **meant to be a guide** that describes some of the relevant dimensions and characteristics and reviews measurement approaches.
1. Data-enhanced and data-enabled businesses
2. The global data value cycle
3. Data monetisation across business models
4. Data types and characteristics
5. Empirical approaches to measurement
80% use data to improve products; 70% see case for new data business; 50% see business case in selling their data

Some examples:

- **Twitter Premium or Enterprise API**
- **Airbus Skywise predictive maintenance**
- **DeepMind medical image scanning**
- **AWS and VW Industrial Cloud connecting machines across plants**
- **New data protection law California (50% threshold)**
The Global Data Value Cycle (1/2)

- Value can arise at any stage of the cycle
- Stages can be in different countries
- Businesses can specialise in certain stages
- Enabled by cloud computing (cheap, sophisticated), IoT, AI
Issue of where to record/report value if aggregated in few data hubs (often transferred between affiliates)

Global Value Chains further fragmented by digitalisation

TeleGeography, McKinsey, 2019
Data monetisation across business models

Revenue source:
- ‘Traditional’
- Selling / licensing of proprietary data
- New products
- Improved products
- More efficient production process

Data-driven revenue (% of total)
- 100%
- 50%
- 0%

Type of business
- e.g. Manufacturers
- e.g. Utility providers
- e.g. Banking or insurance
- e.g. Online platforms

Type of internal data used
- Customer data
- IoT (machine sensors)
- Production errors
- Energy usage patterns
- Payment methods
- Credit rating data
- Fraud data
- Payment patterns
- User clickstreams, location, language, browser type, device, tastes

Type of external data used
- Prices of raw materials, energy
- Weather and climate
- Demographic data
- Energy prices, exports and imports
- Real-time grid reliability
- Interest rates
- Macroeconomic variables (inflation, GDP)
- Consumer surveys
- Internet penetration rates
- Smartphone adoption rates

Example of data monetisation
- Real-time turbine monitoring
- Peak price surcharge
- Automated fraud detection software
- Targeted advertising

‘Data-enabled’ business

‘Data-enhanced’ business
Types of data

- **Private sector data**
- **Public sector data**
- **Proprietary data**
- **Open/public domain data**
- **Personal data**
- **Organisational data**

- **Data source**
- **Data generation**
- **Data subject**
- **Funding of data collection & maintenance**

**To understand:**
- Internal data
- External data
- User created data
- Machine generated data
## Data characteristics

Linkable, Accessible, Disaggregated, Timely, Trustworthy, Representative, Scarce

<table>
<thead>
<tr>
<th>Data characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linkable</strong></td>
<td>Ability to (easily) merge with the other datasets</td>
</tr>
<tr>
<td><strong>Accessible</strong></td>
<td>Ability to (easily) retrieve and/or integrate into business processes</td>
</tr>
<tr>
<td><strong>Disaggregated</strong></td>
<td>Availability at the desired level of aggregation</td>
</tr>
<tr>
<td><strong>Timely</strong></td>
<td>Updated with sufficient frequency to meet the business requirements (annually, daily or in real time)</td>
</tr>
<tr>
<td><strong>Trustworthy</strong></td>
<td>Deemed credible by users; unbiased and impartial, and not dependent on judgment, interpretation, or evaluation of individuals</td>
</tr>
<tr>
<td><strong>Representative</strong></td>
<td>Records do not contain missing fields and are representative enough to meet business requirements</td>
</tr>
<tr>
<td><strong>Scarce</strong></td>
<td>Data are proprietary or secret, difficult to come by</td>
</tr>
</tbody>
</table>
1. Market-based approaches

- Challenge: no well-defined market and infrequent transactions
- Value is context-dependent. WTA of individuals to reveal social security number is US$ 240 but it can be bought for US$ 15 on the market
- Data brokers such as LexisNexis or Experian collect, aggregate, store and sell information on individuals. Pricing models unknown.
- Opportunity: explore data-seeking M&A activities or insurance market (e.g. against data breaches)
2. Data as knowledge-based asset

- Insights and measurement approaches from intangibles literature
- Databases qualify as asset under SNA 2008, but ”data” explicitly excluded.
- Databases only include management systems and cost of digitisation (Ahmad & van de Ven, 2018).
- Value of data in firm acquisitions is goodwill and database acquisitions it is GFCF. Data as intermediate input invisible in accounts?
- Experimental: Explore bankruptcy cases where value of data has been negotiated in court (e.g. RadioShack in 2015)
3. Valuation based on business models and data value cycle

- Select key businesses for different business models and analyse data monetisation at different stages of the value cycle
- Use similar methods to capitalisation of R&D to infer value of data
- Li et al. (2019) estimate value of Amazon Marketplace data is US$125 billion (16% of market cap)
4. Use input-output tables to track data flows

- Data as intermediate good
- Challenge: need to assign ”data intensity” of sector outputs and assume depreciation rates
- Opportunity: Use information job openings in data-related jobs. Statistics Canada recently published some assumptions on ”time spent producing data” by occupation
- For Canada this method yields annual investment in data in range of 7-11 billion USD
Can we see data everywhere but in business balance sheets?

1. Our economies are increasingly powered by data. Consistent and comparable measures are needed for policy design.
2. Data is a special intangible asset: its value is highly context- and characteristics-dependent, costs of reproduction close to zero, and moving data around globally is easy, fast, and cheap (cloud computing as GPT). We need global approach to measurement.
3. As first step, statistical offices should introduce new questions in surveys.
4. Much more research needed on how businesses monetise data across different business models.
Thank you for listening!

Get in touch for the discussion paper:
d.nguyen@niesr.ac.uk