Measuring Data in the National Accounts

Robert Kornfeld

BEA Advisory Committee Meeting

May 10, 2019

Changing Role of Data: IoT Sensor Data

- Caterpillar and John Deere install sensors on machinery and equipment to collect data on operations.
- Sensor data are combined with historical and real-time data on weather, irrigation, and other relevant systems.
- Actionable intelligence is available in platforms to reduce downtime, save fuel, improve safety, and manage crops and worksites.

Changing Role of Data: Personal Data

- Insurance firms collect health and lifestyle data from fitness devices and social network sites to assess risk.
- Location firms collect geolocation data from mobile service providers and mobile apps to sell to advertisers, retailers, and investment firms.
- E-commerce firms collect purchase and browsing data to predict demand and improve service.

Stats Canada-BEA Collaboration on Data

- What is the role of data in a modern economy?
- What is an appropriate typology of data?
- How are data currently treated in national accounts and how are data valued in the private and public sectors?
- Who owns data?
- What are the different methods that national statisticians could use to assign a value to data?
- What is the value of data in Canada and the U.S.?

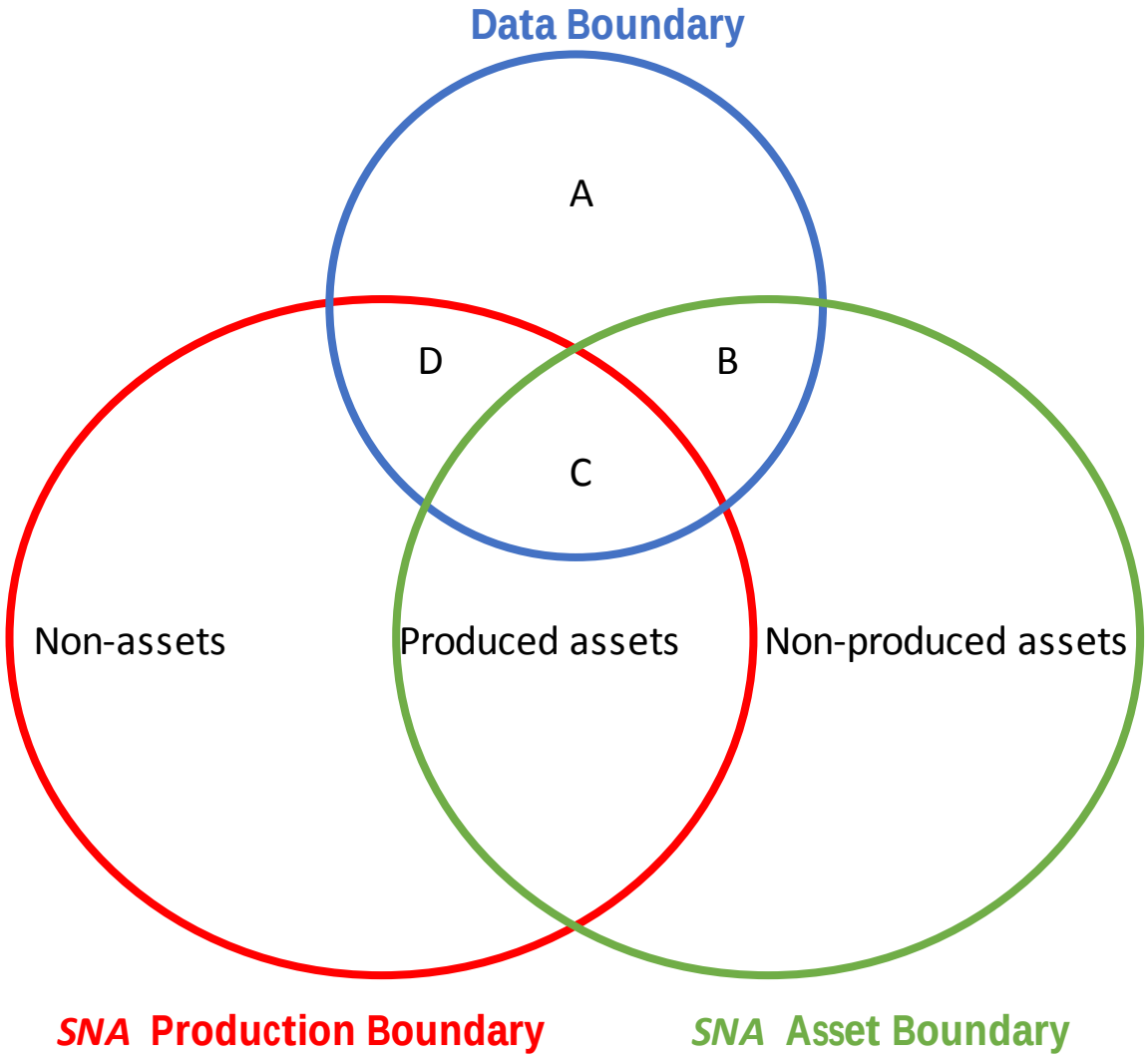
SNA Recommendations on Data

- Databases are within scope of the *SNA* asset boundary
  - Exclude value of data in own-account databases
  - Include value of data in market purchases of databases
- Data as capital formation
  - Canberra II Group focused on data as a knowledge asset
  - Is data a knowledge asset or an information asset?
- No guidance on data as intermediate consumption
  - May be exchanged in traditional B2B transactions
  - May be exchanged in non-traditional C2B transactions

Data in the U.S. National Accounts

- Consistent with *SNA* recommendations
  - No distinction between software and databases
- Purchased software
  - Based on receipts data from Economic Census or SAS
  - NAICS: software publishers (5112), data processing and hosting (5182), computer systems design (5415)
- Own-account software
  - Based on wage data from Occupational Employment Statistics times a blow-up factor to reflect all operating costs
  - OES: programmers (15-1131), app developers (15-1132), system developers (15-1133), systems analysts (15-1121)

Considerations for Data: Data Boundary



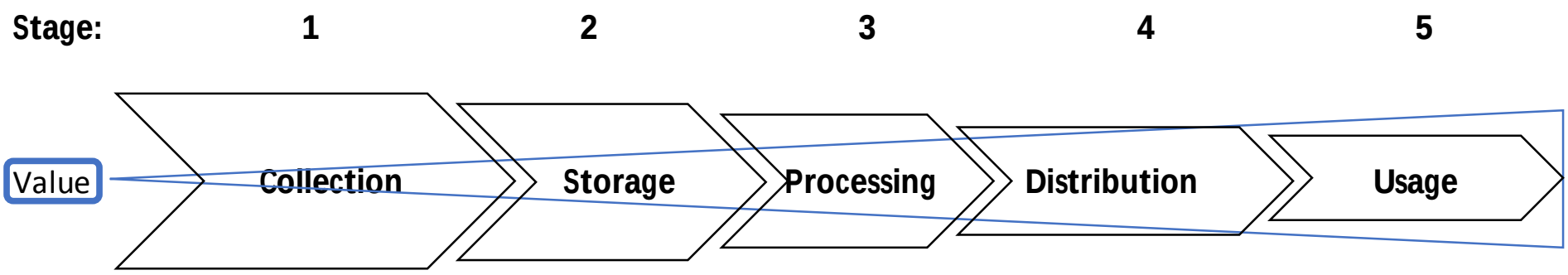
Source: Adapted from *SNA 2008* .

Considerations for Data: Classification

Personal Data	*Institutional Data*		
	Businesses	*Governments*	*Non-Profits*
User-Generated	Personnel Files	Personnel Files	Personnel Files
Behavior	Accounting Records	Accounting Records	Accounting Records
Social	Legal Docs	Legal Docs	Legal Docs
Location	Financial Docs	Financial Docs	Financial Docs
Demographic	Customer Lists	Intelligence Records	Social Policy Programs
Official Identification	IoT Sensors	Diplomatic Cables	Public Policy Programs
		Defense Files	
		Statistical Surveys	
		Regulatory Records	
		Admin Records	
		Monitoring Tech	

Source: Adapted from World Economic Forum (2011) and OECD (2013).

Considerations for Data: Data Value Chain



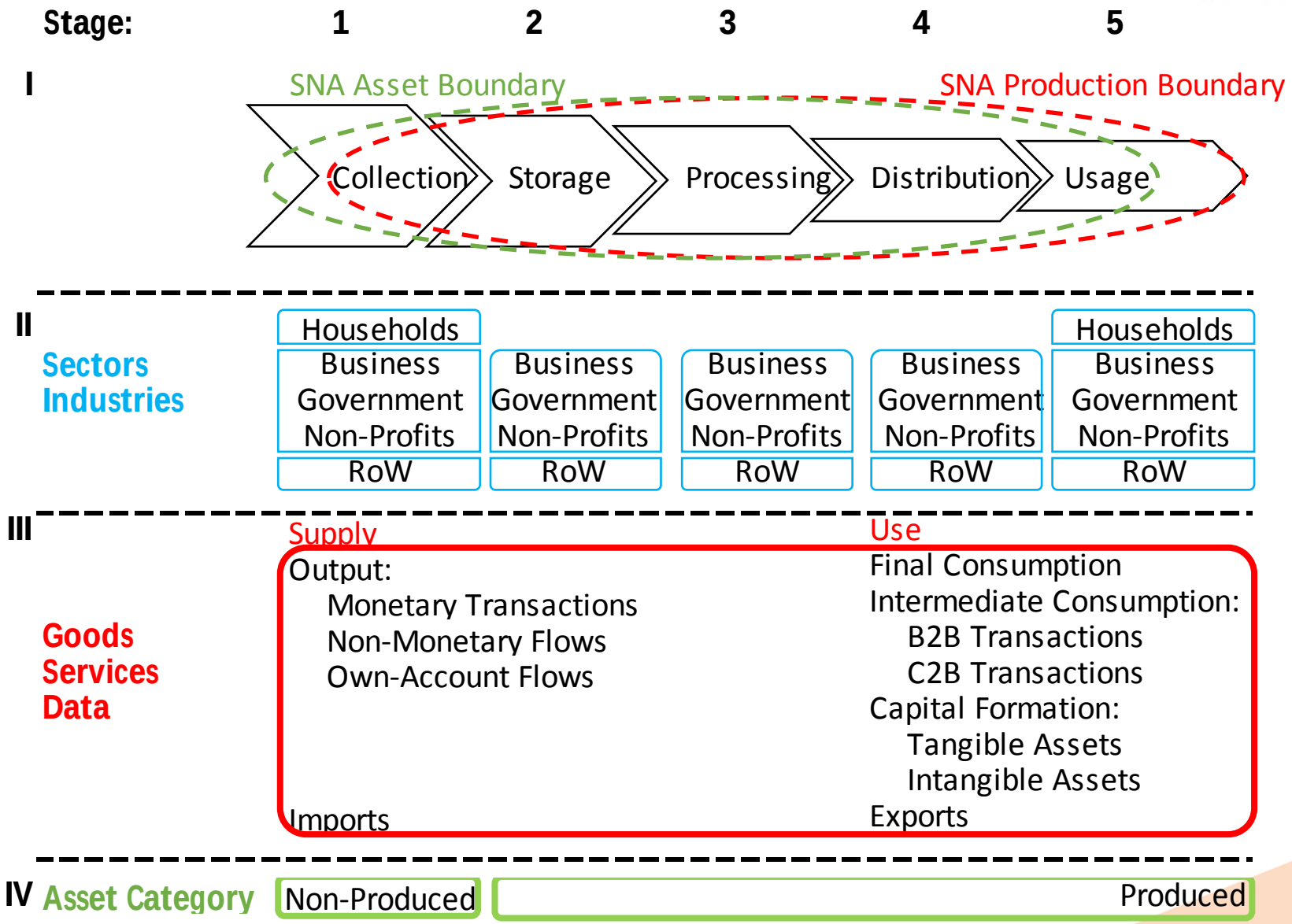
Collection	Storage	Processing	Distribution	Usage
Electronic payments	Cloud facilities	Analysis	Tracking companies	Business models
Web-connected devices	Data centers	Fusion	Data exchanges	Regulatory programs
Regulatory filings				Social programs
				Information
				Entertainment

Source: Adapted from OECD (2013) and Moro Visconti et al. (2017).

Considerations for Data: Features

- Features shared with other *SNA* assets
  - Non-rival but excludable: risk of multiple counting
  - Irregular depreciation profiles: may experience appreciation (Li et al. 2019)
- Features unique to data
  - Characteristics of both goods and services (Mandel 2012)
  - Provision of personal data by households at beginning of the value chain is not customary in national accounts
  - Value of data is context dependent (OECD 2013)

Data in the *SNA* Framework



Valuation Methods

- National accounts
  - Market-based: observed transactions
  - Cost-based: sum of costs
  - Income-based: discounted present value of future profits
- OECD (2013)
  - Market prices
  - Market capitalization or revenue(net income) per data record
  - Economic costs of identity theft or data breaches
  - Insurance premiums for identity protection
  - Experiments and surveys to assess willingness to pay

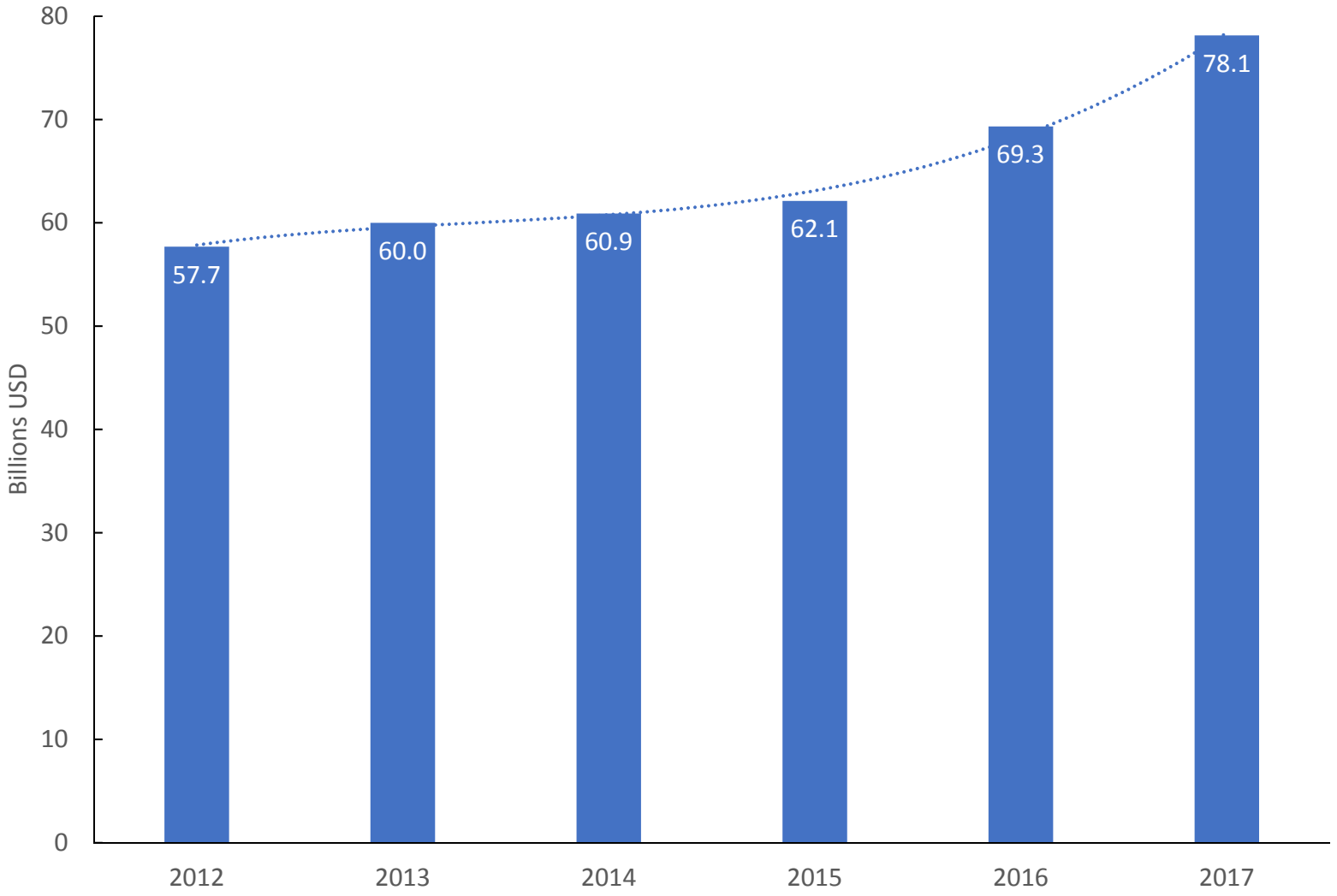
Valuing Data for the U.S. Economy

- Cursory measures establish some context
  - Do not provide a complete picture
  - Allow assessment of currently available sources and methods
  - Provide insight into gaps in classification
- Limit measures to available official statistical sources
  - Economic Census (EC)
  - Service Annual Survey (SAS)
  - Occupational Employment Statistics (OES)
- Assume zero value for government and non-profit data

Purchased Data-Related Products

- Receipts from 2012 EC and SAS for 2013-2017
- NAICS 518 (data processing and hosting)
- Direct measure: Four products from EC
  - Data storage services
  - Data management services
  - Information and document transformation services
  - Other data processing or IT infrastructure provisioning
- Indirect measure: Twelve EC products grouped in one SAS product

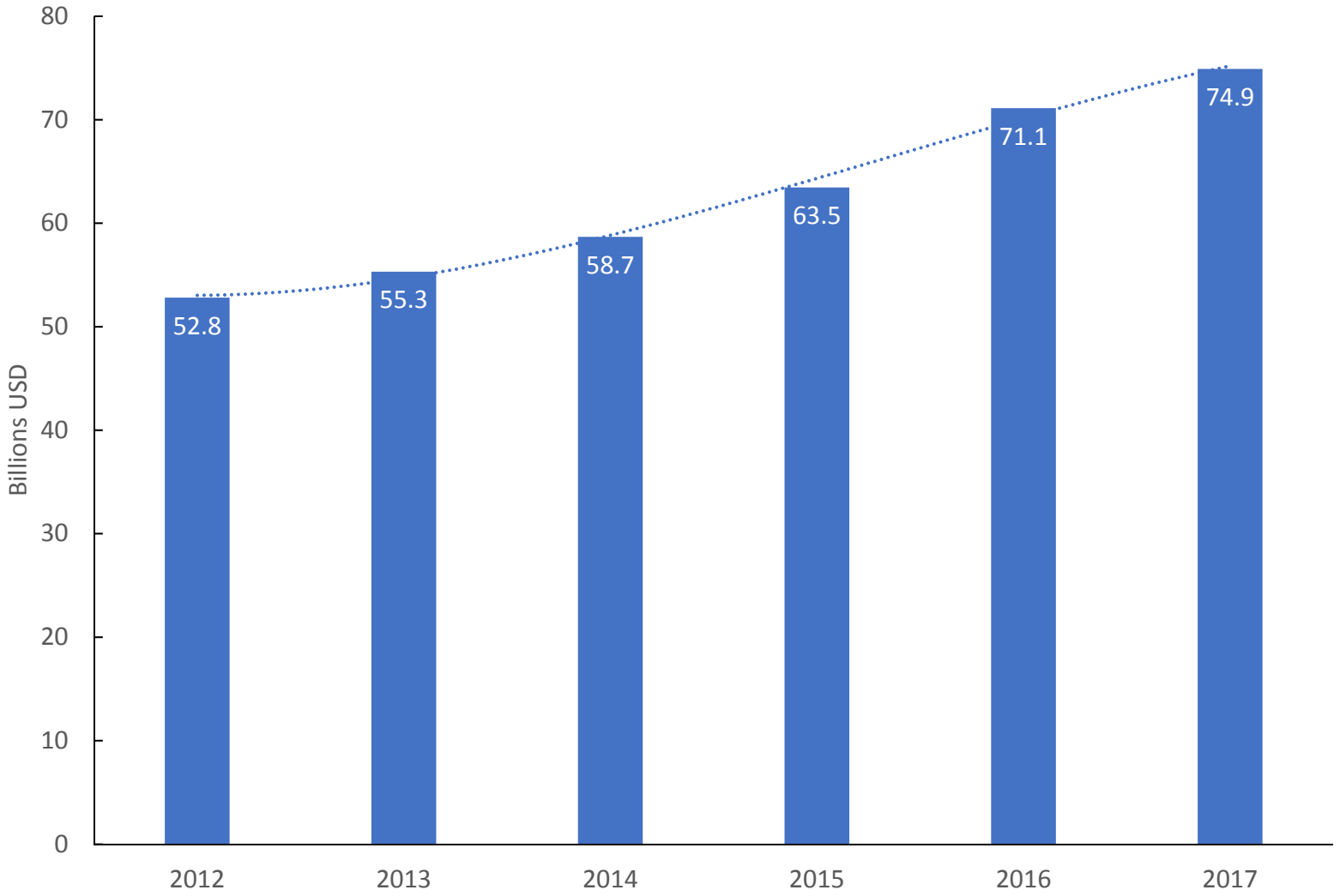
Indirect Data-Related Product Receipts for NAICS 518



Own-Account Data-Related Products

- Wages from OES for 2012-2017
- Seven occupations for computers and mathematics
  - Information security analysts (15-1122)
  - Database administrators (15-1141)
  - Actuaries (15-2011)
  - Mathematicians (15-2021)
  - Operations research analysts (15-2031)
  - Statisticians (15-2041)
  - Mathematical technicians (15-2091)
- Multiply total wages by a blow-up factor for other costs

Data-Related Production Costs for Private Industries (except NAICS 518)



Conclusions

- A classification for data is necessary to determine what data are in and out of scope of the *SNA* boundaries.
- Treatment of data in national accounts faces some of the same challenges as capital measures of R&D, software, and other assets.
- The cursory estimates we present demonstrate that measures of data-related products such as data mining and management are possible with existing classifications, statistical sources, and methodologies.