Outcome of the global consultation on

Testing - DZ.6 Recording of Data in the National Accounts

CONSENTERS ONLY

A total of 19 respondents contributed to this consultation, 14 of which consented to publishing their verbatim responses which are provided below. <u>However, the tables reflect the answers of all 19 respondents.</u>

Completely anonymous contributions are excluded.

1A. Do you have data sources allowing estimating the value of own-produced data assets according to those cost components?

Q1A	Frequency
Yes, partly	6
No, but may be possible in a few years	4
No, and no development is in sight in a foreseeable future	9
Total	19

Australia (Australian Bureau of Statistics): Yes, partly

Canada (Statistics Canada): No, and no development is in sight in a foreseeable future

Chile (Central Bank of Chile): No, but may be possible in a few years

Denmark (Statistics Denmark): Yes, partly

Finland (Statistics Finland): No, and no development is in sight in a foreseeable future

Germany (Federal Statistical Office (Destatis)): Yes, partly

Netherlands (Statistics Netherlands): No, and no development is in sight in a foreseeable future

Norway (Statistics Norway): No, and no development is in sight in a foreseeable future

Qatar (Planning and Statistics Authority): No, and no development is in sight in a foreseeable future

Republique de Guinée (Institut National de la Statistique): No, and no development is in sight in a foreseeable future

Slovenia (Statistical Office of the Republic of Slovenia): No, but may be possible in a few years

Spain (INE - NATIONAL STATISTICS OFFICE): No, but may be possible in a few years

Sudan (Central Bureau of statistics): No, and no development is in sight in a foreseeable future

US (BEA): We have used online job posting data to estimate time use factors per occupation as an input to out sum of cost approach to obtain an effective payroll. However, these estimates reflect the aggregate time-use across the various activity components and have not attempted to disaggregate costs by each of the components.

1A. Please elaborate 1B. What is the frequency of these data sources?

Australia (Australian Bureau of Statistics): The ABS has census and labour force occupations data available that allows for the estimation of own account produced assets. We currently have no specific data available to inform the non-labour cost components. Census information every 5 years, and labour force data quarterly.

Chile (Central Bank of Chile): It is planned to adjust business surveys to collect these kinds of costs. Currently we count with transactional administrative records of all contributors to identify purchases of data (electronic invoices).
Business surveys are collected annually and electronic invoices are available in a daily frequency

Denmark (Statistics Denmark): Data on specific occupations by industry like for software. It must be investigated in more detail what can be used and if there is potential overlaps. Annual

Finland (Statistics Finland): New surveys or the widening of the existing ones.

Germany (Federal Statistical Office (Destatis)): Results from the Micro Census provide information on employment numbers according to occupational groups and economic sectors. The Earnings Survey provide information on annual average earnings by occupational groups. After making challenging assumptions
The Micro Census is published annually. Results from the structural earning surveys are published every 4 years.

Netherlands (Statistics Netherlands): You either need more information on the costs on the business side. This would require an expansion of the SBS. Alternatively you would need yearly information on what proportion of time occupations spend on the mentioned activities.

Norway (Statistics Norway): Better and more refined data on how much time relevant professions actually spend on the production of data.

Qatar (Planning and Statistics Authority): Data collection tools will need to be either adjusted or developed.

Slovenia (Statistical Office of the Republic of Slovenia): The idea is to use/expand the survey on ICT usage in enterprises (to obtain shares of staff time spent on data production), NA GDP production and income approach data and GFCF survey (covering purchases/sales of databases). NOS estimate may be challenging
The frequency of all sources is annual.

Spain (INE - NATIONAL STATISTICS OFFICE): intermediate consumption and investment Survey 2024 - this is an specific survey performed by National Accounts Department in INE, not required in ESA regulation. We are planning to insert a new module covering all the required inputs for new Data asset.
Plan A - annual: try to compile estimates mixing LFS and Labour Cost Survey for specific ocupations and industries in Data own-final use.
Plan B - every 4 years: Intermediate Consumption and Investment Survey - try to insert a whole new module with questions about data, artificial intelligence, market brands+naming and all new assets planned in SNA 2025.

US (BEA): The number of employees by industry of privately held establishments as well as average wages and salaries is obtained from the U.S. Bureau of Labor Statistics OEWS program and values are available at an annual basis. The markup factor from payroll to total production cost is usually a multi-year average based on annual estimates. The estimated time-use factors are estimated using data from multiple years and currently the research/experimental estimates use time-invariant estimates.

1C. Please describe the methods you (may) apply to your data sources to estimate the value of own produced data assets according to the cost components above.
Australia (Australian Bureau of Statistics): To be provided with updated data estimates.

Chile (Central Bank of Chile): In the case of business surveys, we would use the structure of the costs from a sample of firms complemented with administrative record to extrapolate them to the population.

Denmark (Statistics Denmark): Most likely the same method as for software and use assumptions when necessary.

Germany (Federal Statistical Office (Destatis)): To estimate investments in own-account data, occupational groups involved in the creation of data have to be selected and the respective involvement rate estimated. Selection and estimation are based on job descriptions provided by the German Occupational Classification (KldB 2010). To take other labour costs, intermediate consumption of fixed capital and a profit margin into account, a general mark-up factor is introduced. These cost components cannot be measured individually with the available data sources.

Slovenia (Statistical Office of the Republic of Slovenia): ICT usage survey (ICT ENT) is currently already used for estimation of own-account software and database development. The survey asks about the number of developers and their average share of time dedicated to development of software and databases for own use by the unit. The next step would be the expansion of the ICT survey questionnaire to try to identify the average share of time spent for producing data (i.e. broader definition than databases) for the other relevant codes of the ISCO classification of occupations (besides IT experts) (though this would probably not be enough to cover the non-ICT profiles involved in data production). An estimate of the own-account production of data and databases could then be obtained by sum-of-cost approach with relevant shares of staff time spent by occupation. The main data source for the production account for most industries are the annual accounting statements.

Spain (INE - NATIONAL STATISTICS OFFICE): Intermediate Consumption and Investment Survey - try to insert a whole new module with questions about data, artificial intelligence, market brands+naming and all new assets planned in SNA 2025. Prepare speciific questions about costs ccomponents and mix this input with SBS microdata to compile the required estimates.

US (BEA): If needed, we could explore various ways to estimate time-use factors per occupations for the various components.

1D. Do you have other comments on this feature?

Australia (Australian Bureau of Statistics): The ABS has census and labour force occupations data available that allows for the estimation of own account produced assets. We currently have no specific data available to inform the non-labour cost components. Similar to that currently used to compile R&D and Software estimates, we will likely need to add/alter questions on our annual economic activity survey or create and run a specific targeted survey to be able to collect specific data asset information. There are currently no plans to do either of these, nor has there been any work identifying what information would need to be collected.

Canada (Statistics Canada): Using a Sum of Costs approach, it would be very difficult separating the costs associated with these three activities.

Denmark (Statistics Denmark): No

Netherlands (Statistics Netherlands): We completed one research project where data was estimated using an occupation approach. However, the fractions used there would have to be updated on a regular basis if this type of estimation was to be used in production. Only one set of fractions was used based on data from other countries. This research will most likely not be repeated on a regular basis.

Norway (Statistics Norway): In our view, this would not contribute to a better estimate of data as capital in the national accounts, as we currently do not possess detailed data on this.

Qatar (Planning and Statistics Authority): Assistance to modify current data collection tools would be needed.

Republique de Guinée (Institut National de la Statistique): pas du tout / <mark>Not at all</mark>

Slovenia (Statistical Office of the Republic of Slovenia): In order to estimate staff time costs, the relevant "specific occupations" should be defined (possibly harmonised on international level). The feasibility of obtaining time-shares spent on "data production" with the ICT survey for non-ICT profiles is also not yet clear. It is paramount to have a well established standardized data source with broad international implementation as a starting point for estimation of data production, such as was the case with the Frascati survey data for the capitalization of R&D.

US (BEA): In terms of the activities contributing to the value of data assets, it would be helpful to have clarity on the scope of those activities with emphasis on two potential activities: (1) data collection and (2) data analysis. In terms of the data collection, a lot of data is collected though social media and "free goods" which provide an online platform design to collect the data by providing free user experience. How should the investment in these platforms be accounted for especially when those might fall under other assets like software? In terms of data analysis, is employing data assets to obtain estimates or training prediction models within scope? Current approach and estimates available at: [Valuing the U.S. Data Economy Using Machine Learning and Online Job Postings | U.S. Bureau of Economic Analysis (BEA)](#).

2A. Do your data sources allow to separately estimate the value of 'computer software' from 'data and databases'?

Q2A	Frequency
Yes, but with some overlap	6
No, but may be possible in a few years	7
No, and no development is in sight in a foreseeable future	6
Total	19

Australia (Australian Bureau of Statistics): No, but may be possible in a few years

Canada (Statistics Canada): Yes, but with some overlap

Chile (Central Bank of Chile): No, but may be possible in a few years

Denmark (Statistics Denmark): Yes, but with some overlap

Finland (Statistics Finland): No, and no development is in sight in a foreseeable future

Germany (Federal Statistical Office (Destatis)): No, and no development is in sight in a foreseeable future

Netherlands (Statistics Netherlands): No, and no development is in sight in a foreseeable future

Norway (Statistics Norway): No, but may be possible in a few years

Qatar (Planning and Statistics Authority): No, and no development is in sight in a foreseeable future

Republique de Guinée (Institut National de la Statistique): Yes, but with some overlap

Slovenia (Statistical Office of the Republic of Slovenia): Yes, but with some overlap

Spain (INE - NATIONAL STATISTICS OFFICE): No, but may be possible in a few years

Sudan (Central Bureau of statistics): No, and no development is in sight in a foreseeable future

US (BEA): Yes, but with some overlap

2A. Please elaborate 2B. Do you have other comments on this feature?
Australia (Australian Bureau of Statistics):
We currently collect information for computer software from our annual economic activity survey, this captures computer software and databases combined. Future work to refine the

questions in this survey may allow us to separate databases and data from computer software. No plans have been made for this yet.

Canada (Statistics Canada): Wages by occupation, assumption is 20% overlap
With refinements, these can be separated and overlap removed.

Chile (Central Bank of Chile): Business surveys and administrative records currently available

Denmark (Statistics Denmark): please see previous answer.
No

Finland (Statistics Finland):
It is very likely that there will be overlapping whatever method is to be used.

Germany (Federal Statistical Office (Destatis)): Basic statistics in Germany must be modified to allow this separation
Separation based on the current statistical method is not feasible

Netherlands (Statistics Netherlands): An expansion of the SBS and investment survey would be required to be able to make this distinction.
At this moment no alternative data sources or research is available to make this distinction.

Norway (Statistics Norway): More refined questionnaires to relevant companies/entitites involved involved in data production.
In our surveys, as of yet, there are no inquiries pinpointing data, but more general inquiries into software investments, and the utilisation of services from IT consultants. However, accommodating for such inquiries could be feasible in the future.

Qatar (Planning and Statistics Authority): Data collection tools will need to be either adjusted or developed.
Assistance would be needed for this topic too.

Republique de Guinée (Institut National de la Statistique):
pas du tout / Not at all

Slovenia (Statistical Office of the Republic of Slovenia): Please see 2B.
Currently the data on own-account computer programming are collected with the yearly survey on ICT usage and e-commerce in enterprises (ICT ENT). The survey asks about the number of developers and their average share of time dedicated to development of software and databases for own use by the unit. The reported amount of time spent in developing the software is separated from the amount spent on developing databases, so the two components could in theory be separated. Acknowledging that in practise the work of IT developers may in parallel include software as well as database development, the usage of more detailed shares may be less reliable and still produce an overlap i.e. in the sense of proper asset type classification, not as double counting. The purchases of software and purchases of databases are also separately reported in the GFCF survey.

Spain (INE - NATIONAL STATISTICS OFFICE): Intermediate Consumption and Investment Survey - try to insert a whole new module with questions about data, artificial intelligence, market brands+naming and all new assets planned in SNA 2025. Prepare speciific questions and mix with SBS microdata
A first estimate could be compiled using ocupations info from LFS by industry.

US (BEA): Computer software is captured in six accounts: (1) federal defense, (2) federal non-defense, (3) state and local, (4) pre-packaged software, (5) custom software, and (6) own-account software. Purchased data is technically included (mostly) in pre-packaged software based on economic transaction data. However, the taxonomy for purchased datasets is less detailed than software leading to estimates being less reliable that for pre-packaged or custom software for which the transaction data is of high quality. For own-account data and databases, we currently do not capitalize those in the accounts but have experimental estimates from a sum of cost approach for data and databases. Once difference between the treatment of software and current approaches for databases is that software transaction data relates almost exclusively on sold copies while the originals are captured in the own account while for data/databases we have not decided on how to account for original vs copies. These estimates are based on occupational time-use factors and exclude occupations engaged primarily on software development to avoid overlap between the two activities/asset classes

We have considered trained models (e.g., machine learning / AI) and data/databases as potentially having various relations. For example, a trained model can work as a substitute or alternative or enhanced representation of the "content" of a data asset (e.g., database). Some alternatives to account for these relations include: (1) treat data and models as two separate capital assets (e.g., bulldozers are used to construct buildings); (2) treat data as embedded capital (e.g., air conditioning unit which is permanently installed in a building is considered part of the building rather than its own capital), or (3) treat data and models as just different terms for the same capital asset. Each alternative has implications for service life, depreciation, and prices.

3A. Do your data sources that allow estimating the value of 'Data assets' used in production excluding 'ancillary data' that are not directly used in production? 3B. Do you think that in a few years you will have data sources allowing to estimate the value of 'Data assets' used in production excluding 'ancillary data' that are not directly used in production?

Q3A	Frequency
Yes	1
No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities	15
No response	3
Total	19

Q3B	Frequency
Yes	2
No	14
No response	3
Total	19

Australia (Australian Bureau of Statistics): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Canada (Statistics Canada): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Chile (Central Bank of Chile): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
Yes: bussines surveys

Denmark (Statistics Denmark): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Finland (Statistics Finland):
No

Germany (Federal Statistical Office (Destatis)): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Netherlands (Statistics Netherlands):
No

Norway (Statistics Norway): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Qatar (Planning and Statistics Authority): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for

productive activities
No

Republique de Guinée (Institut National de la Statistique): Yes

Slovenia (Statistical Office of the Republic of Slovenia): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Spain (INE - NATIONAL STATISTICS OFFICE): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

Sudan (Central Bureau of statistics): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities
No

US (BEA): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities

No, It may require instruments such as surveys that can provide some estimates of how various industries are using data assets in production.

3C. Do you have or will have data sources allowing to separately estimate data as an asset excluding ancillary data specifically for non-market producers?

Q3C	Frequency
Yes	2
No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities	2
No response	15
Total	19

Chile (Central Bank of Chile): No, but an assumption can be made that occupations included in the sum-of-cost calculation are only producing data used for productive activities

Republique de Guinée (Institut National de la Statistique): Yes

US (BEA): Potentially if we were able to have access to additional data sources for these estimates such as surveys.

3D. Do you have other comments on this feature?

Australia (Australian Bureau of Statistics): It will be difficult, if not impossible, to collect information required to separately identify ancillary data and data which is used directly in production. It will likely also be difficult to make assumptions on which occupations to included that will only include data produced for productive activities, data is created and used by many different occupations and the occupations themselves does not give sufficient insight into the productiveness of the data.

Denmark (Statistics Denmark): No

Finland (Statistics Finland): We do not have direct data sources concerning data so the separately estimates are not possible.

Norway (Statistics Norway): The already stated challenges with measuring data as a separate asset class will also be relevant with regards to ancillary data. If measuring data as a separate asset class from databases and software is difficult due to a lack of relevant data, a further detailed estimation of subcategories of data would be difficult.

Qatar (Planning and Statistics Authority): No.

Republique de Guinée (Institut National de la Statistique): pas du tout / <mark>Not at all</mark>

Slovenia (Statistical Office of the Republic of Slovenia): The definitions of data in the survey questionnaires instructions can in theory exclude ancillary data.

US (BEA): The asset boundary definition should include illustrative examples that national statistical offices could use to best implement the guidance. For example, are payroll, transaction data, and employee performance evaluations a dataset used in production? One could argue that the data is vital for businesses to operate as well as potentially being valuable for policy, research (e.g., design of hiring, retention packages and policies), and required for sustaining other types of "assets" such as human or organizational capital. However, many of these datasets would have to be collected for regulatory or compliance purposes even if a company does not actively use these beyond their most limited role.

4A. Do you have information on services lives of data and databases? 4B. What is your source of information on service lives?

Q4A	Frequency
Yes	2
No information is available at the moment	16
No response	1
Total	19

Australia (Australian Bureau of Statistics): Yes
Other research activities you carried out

Canada (Statistics Canada): No information is available at the moment

Chile (Central Bank of Chile): No information is available at the moment

Denmark (Statistics Denmark): No information is available at the moment

Finland (Statistics Finland): No information is available at the moment

Germany (Federal Statistical Office (Destatis)): No information is available at the moment

Netherlands (Statistics Netherlands): No information is available at the moment

Norway (Statistics Norway): No information is available at the moment

Qatar (Planning and Statistics Authority): No information is available at the moment

Republique de Guinée (Institut National de la Statistique): Yes
Business survey

Slovenia (Statistical Office of the Republic of Slovenia): No information is available at the moment

Spain (INE - NATIONAL STATISTICS OFFICE): No information is available at the moment

Sudan (Central Bureau of statistics): No information is available at the moment

US (BEA): No information is available at the moment

5A. Do you have information on retirement and depreciation functions of data and databases?
5B. What is your source of information on retirement and depreciation functions?

Q5A	Frequency
Yes	1
No information is available at the moment	17
No response	1
Total	19

Australia (Australian Bureau of Statistics): No information is available at the moment

Canada (Statistics Canada): No information is available at the moment

Chile (Central Bank of Chile): No information is available at the moment

Denmark (Statistics Denmark): No information is available at the moment

Finland (Statistics Finland): No information is available at the moment

Germany (Federal Statistical Office (Destatis)): No information is available at the moment

Netherlands (Statistics Netherlands): No information is available at the moment

Norway (Statistics Norway): No information is available at the moment

Qatar (Planning and Statistics Authority): No information is available at the moment

Republique de Guinée (Institut National de la Statistique): No information is available at the moment

Slovenia (Statistical Office of the Republic of Slovenia): No information is available at the moment

Spain (INE - NATIONAL STATISTICS OFFICE): No information is available at the moment

Sudan (Central Bureau of statistics): No information is available at the moment

US (BEA): Yes, On very limited cases, there are laws determining the service life of certain data assets. For example, personal information used for credit history has a set limit after which it must be discarded and not included. For data assets in general, we are currently using a placeholder of software.

6. Do you have other comments on this feature?

Australia (Australian Bureau of Statistics): In response to Q4A:
Limited but more than none. We do not currently have direct information on the asset life of data and databases. During the production of our experimental estimates on data, we worked with the assumption that most data is used for the purpose of understanding behaviour and that most of the commercial value that is extracted from data is linked to economic behaviour and consumer preferences.

To understand the useful life of data related to consumer preferences we looked to work published in the following information paper: 6401.0.60.002 - Information Paper: Increasing the Frequency of CPI Expenditure Class Weight Updates, July 2016 (abs.gov.au)

To reduce the bias in the CPI arising from consumers substituting between different combinations of products, the frequency of reweighting the CPI was increased. This change was supported by the empirical evidence in the paper and suggested that an appropriate mean asset life for data could be of 3 years and a maximum asset life of 5 years.

These assumptions and evidence are further supported by the DZ6 guidance note, which suggests the use of a shorter asset life to account for the inclusion of all own account produced data in the asset boundary.

Canada (Statistics Canada): we would likely borrow rates from software for these categories.

Denmark (Statistics Denmark): No

Netherlands (Statistics Netherlands): We do have information on parameters for PIM for software and databases but not for data assets separately.

Norway (Statistics Norway): Currently, we have no information on service life of data and databases.

Qatar (Planning and Statistics Authority): No.

Republique de Guinée (Institut National de la Statistique): pas du tout / <mark>Not at all</mark>

Slovenia (Statistical Office of the Republic of Slovenia): We should strive for harmonisation of PIM parameters between the countries.

US (BEA): While engaging in talks with other national statistical offices, it has surfaced that it would be beneficial for the measurement community for various offices to specialize on certain aspects to avoid duplicated efforts. For example, research on depreciation or data trade flows may be applicable to various economies. This approach may free resources to further the research on other particulars ahead of the upcoming SNA revision.

Another consideration for the service life and categorizing data assets as long or short-lived depends on how those are defined. For example, streaming data such as latest news, last 24 hours of social media posts, last seven days of temperature have a reasonable definition which gets constantly updated. Are these to be considered short-lived since the underlying values are being updated or long-lived as the asset is defined in terms of the subscription-like service?

7. In order to maximize transparency, we would like to publish responses to global consultations. Do you give consent that your response to this questionnaire can be published?

Q7	Frequency
Yes	14
No	4
No response	1
Total	19

Australia (Australian Bureau of Statistics): Yes

Canada (Statistics Canada): Yes

Chile (Central Bank of Chile): Yes

Denmark (Statistics Denmark): Yes

Finland (Statistics Finland): Yes

Germany (Federal Statistical Office (Destatis)): Yes

Netherlands (Statistics Netherlands): Yes

Norway (Statistics Norway): Yes

Qatar (Planning and Statistics Authority): Yes

Republique de Guinée (Institut National de la Statistique): Yes

Slovenia (Statistical Office of the Republic of Slovenia): Yes

Spain (INE - NATIONAL STATISTICS OFFICE): Yes

Sudan (Central Bureau of statistics): Yes

US (BEA): Yes