STATISTICS

AI, Data, and the Update of the System of National Accounts

NOVEMBER 21, 2024

Erich Strassner
Chief, Real Sector Statistics Division
IMF Statistics Department

AI and Data: Motivation

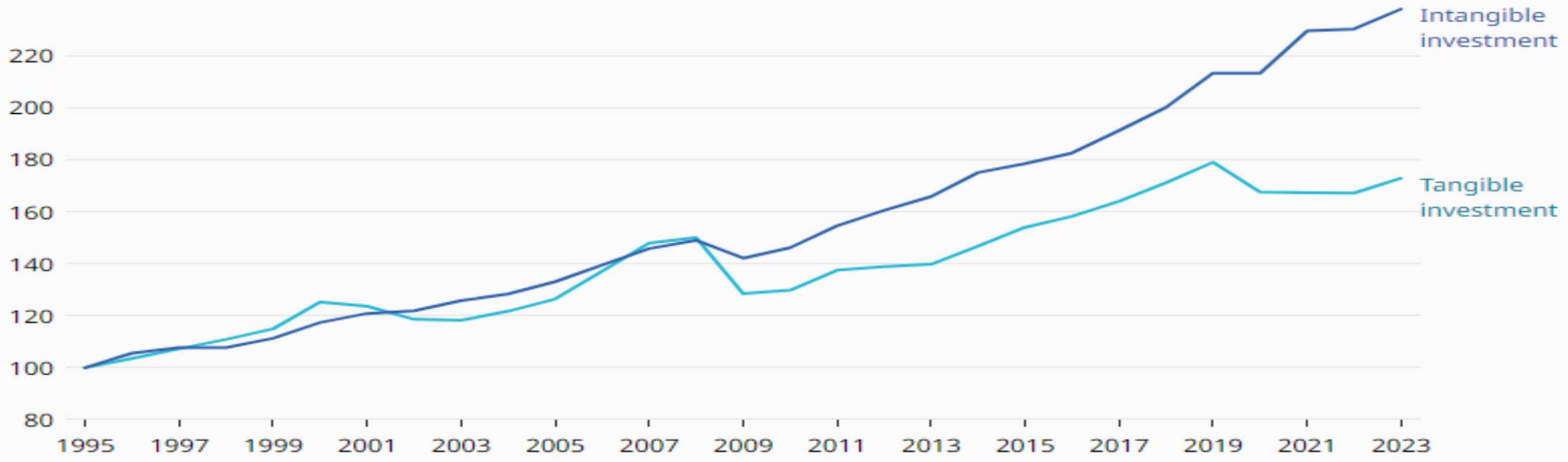
Source: Pixaby

Source: CIO

- Since the 2008 System of National Accounts (SNA) was published, production of data is now fundamental to many business models, while most traditional businesses now produce some form of data to either increase returns or lower costs.
  - A large amount of this data is used repeatedly in production for more than one year, but no fixed asset category currently exists outside of possible inclusion within databases.
- In addition, Generative AI has been progressively developing over the years, but a significant leap in usability and public awareness occurred around 2020 with the release of models like OpenAI's GPT-3.
  - This model demonstrated advanced capabilities in natural language processing, enabling a wide range of applications in various industries.
- Moreover, data is an input into the development of Large Language Models (LLMs). If we do not value data as an asset we may be mis-representing the LLM production function.
- Because of the current classification and definition within the SNA, there has been growing concern that data and AI are not visible and being appropriately represented in the accounts, impacting the estimates of GDP, balance sheets, and productivity.

Investment in Intangible Assets (1)

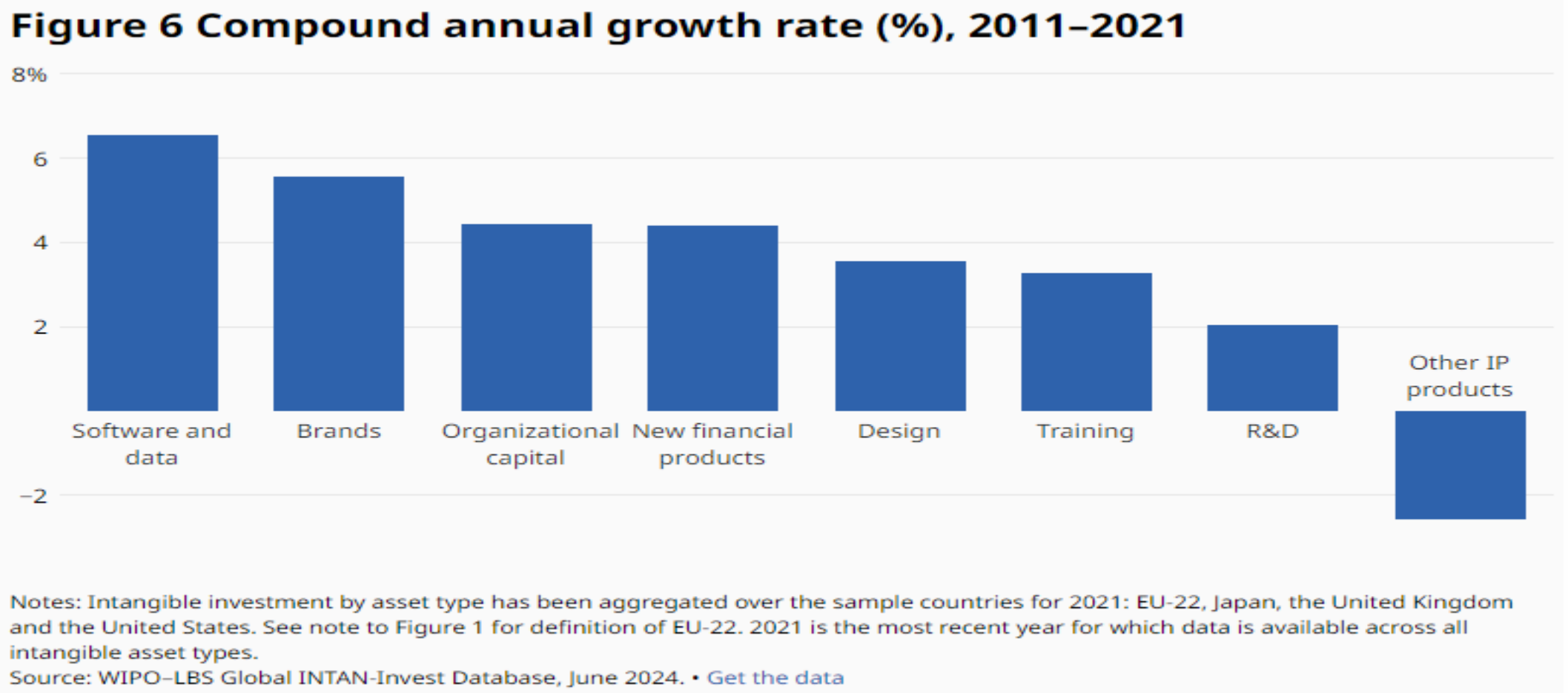
Figure 1 Total intangible and tangible investment, 1995–2023, indexed (1995=100)



Notes: Intangible and tangible investment have been aggregated over the sample countries: EU-22, India, Japan, the United Kingdom and the United States. Estimates are in terms of chain-linked volumes (reference year 2015). The EU-22 economies are Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Latvia, Lithuania, Luxembourg, the Kingdom of the Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain and Sweden.
Source: WIPO–LBS Global INTAN-Invest Database, June 2024. • Get the data

Source: World Intangible Investment Highlights: https://www.wipo.int/edocs/pubdocs/en/wipo-pub-rn2024-32-en-world-intangible-investment-highlights.pdf

Investment in Intangible Assets (2)



Source: World Intangible Investment Highlights: https://www.wipo.int/edocs/pubdocs/en/wipo-pub-rn2024-32-en-world-intangible-investment-highlights.pdf

Key AI Recommendations Endorsed for SNA Update

- Update SNA definition of Intellectual Property Products (2008 SNA 10.98) to:

  The result of research, development, investigation, or innovation leading to knowledge ***or the creation of intelligent systems*** that the developers can market or use to their own benefit in production because use of the knowledge or system is restricted by means of legal or other protection.

- The following definition of AI to be used in the updated SNA:

  "AI is a computer program operating a system capable of recognition, reasoning, communication, and prediction simulating human recognition, reasoning, and communication."

- Give AI explicit visibility within the asset classification by adding an "of which" category under Computer Software. This would permit countries to distinguish between AI and non-AI computer programs when analytically useful to do so.

Key Data Recommendations Endorsed for SNA Update

- Definition of Data: *information content that is produced by **accessing** and **observing** phenomena, **recording,** and **storing** information elements from these phenomena in a **digital format** and that **provides an economic benefit** when used in productive activities*
  - Data is the result of production, i.e., when capitalized in the national accounts, it is classified as a produced asset
  - Data is distinct from the information elements of 'observable phenomena' (OP), which are inputs for data
- Expenditure undertaken to access and record OPs that are added to an established data asset is considered new gross fixed capital formation, as it prolongs the life of the data asset
- Only data that provides an economic benefit when used in the productive activities of its owner for more than one year is included in the SNA asset boundary
- Data is considered part of the intellectual property product suite of assets and will be classified to a newly created asset category called "data and databases"

Economic Measurement: Valuation of AI and Data

- Summary: Create a new category of fixed asset under “Computer Software, Data and Databases”
  - Computer Software “of which” subcategory for AI
  - Data and databases category compiled and presented separately from computer software.
    - In practice it may not always be possible to separate the costs of database structure from that of the database content (i.e. data), thus combined category.
- Valuation of AI and Data assets:
  - Purchased: valued at market prices
  - “Own-account”: valued using cost-based approach

Experimental Estimates of Data Assets and Next Steps

Country	Year	Value of data asset, % of total GDP	PPT difference in total GDP growth for year	PPT difference in total GCF growth for the year
Australia	2016	2.9%	0.016%	0.57%
Canada	2018	1.9%	-0.037%	-0.09%
Netherlands	2017	3.0%	-0.012%	-0.12%
India	2019	1.0%	0.000%	0.14%
USA	2020	0.8%	0.047%	0.26%

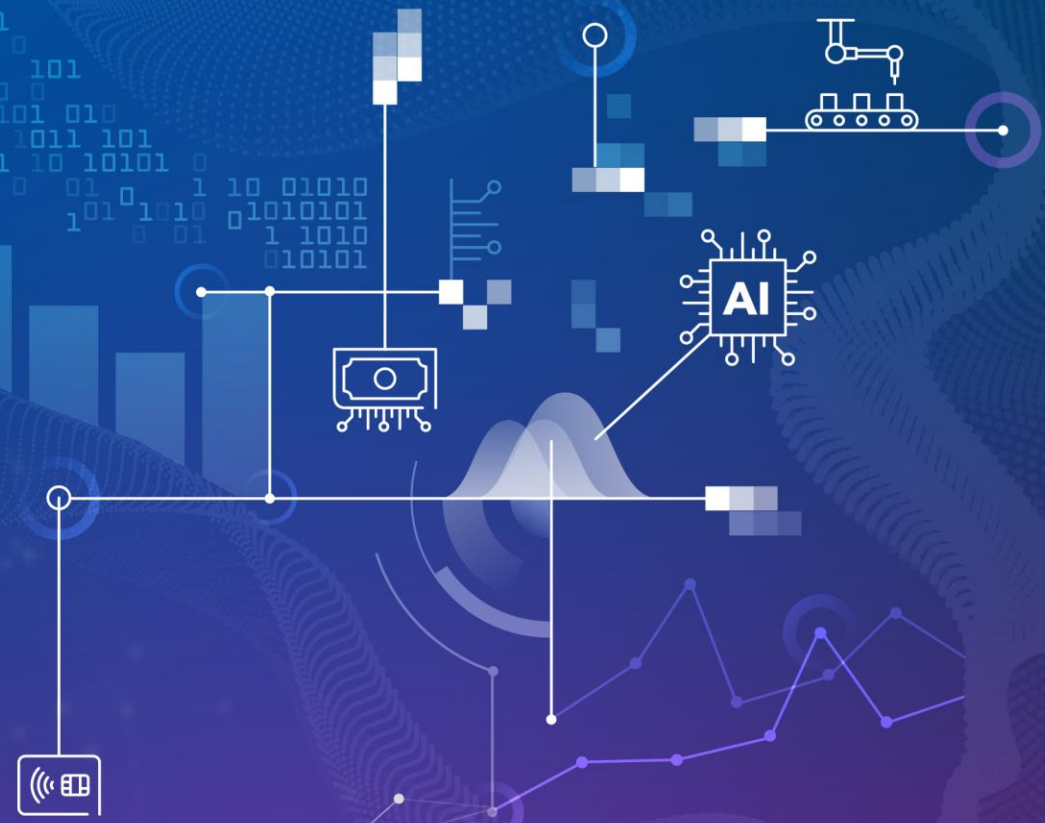
- Currently a Eurostat-IMF Task Force is developing a Handbook on Measuring Data as an Asset, which builds on efforts at the country level.
- Experimentation on the value of AI software in National Accounts is a high priority next step.

12th IMF Statistical Forum

MEASURING THE IMPLICATIONS OF AI ON THE ECONOMY

NOVEMBER 20-21

Washington, DC



#StatsForum