

Research Spotlight

Measuring Open Source Software Innovation

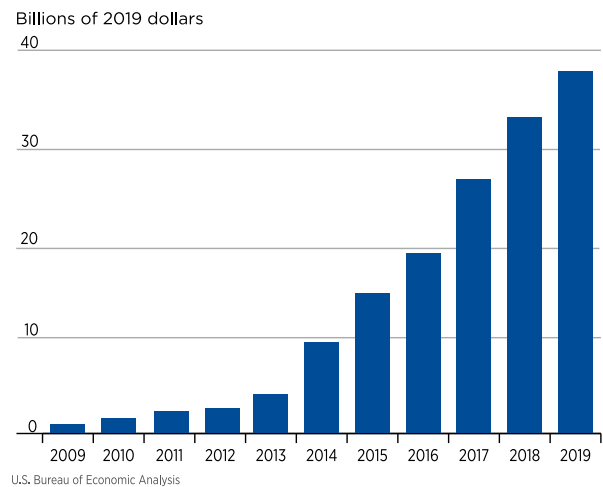
February 20, 2024

Open source software (OSS) is software that anyone can review, modify, and distribute freely, usually with only minor restrictions. Notable examples of OSS include the Linux operating system, Apache server software, and programming languages R and Python. The use and creation of OSS has grown rapidly in recent years due to its contribution in the business sector and the overall economy, but there have been no explicit estimates of investment in this critical public asset.

An effort to fill that measurement need is presented in a recent study by researchers J. Bayoán Santiago Calderón and Ledia Guci of the U.S. Bureau of Economic Analysis (BEA), Gizem Korkmaz of Westat, Brandon L. Kramer of Edge & Node, and Carol A. Robbins of the National Center for Science and Engineering Statistics. The article is available at [the journal Research Policy](#), and an [earlier version](#) was published by BEA as a working paper in July 2022.

BEA regularly publishes statistics of investment in software in the National Income and Product Accounts. These statistics follow the System of National Accounts framework; however, they do not provide a detailed view necessary for understanding differences in proprietary and OSS activity. While some OSS activity is included in BEA's statistics, this activity is not explicitly measured and reported. Explicit measures of investment in OSS are important to understand the relative contributions to OSS by each sector, such as the business sector or the government. For example, household contributions to OSS fall outside the production boundary of the national accounts and are not included in BEA estimates. The featured work offers alternative data decompositions and definitions that provide additional insights on this class of software.

U.S. Real Investment in Open Source Software (2009–2019)



The study presents a framework to measure the value of OSS using data collected from GitHub, the world's largest community where software is developed and shared. The data include over 7.6 million repositories or projects where software development activities are observed. The authors collect information about contributors and development activity such as who made what code changes and when, as well as the licensing of the product. By adopting a cost-estimation model from software engineering, they develop a methodology to generate estimates of investment in OSS that are consistent with the U.S. national accounting methods used for measuring software investment. The methodology draws from the current economic measurement of own-account software, which is software created using internal resources as opposed to purchased or outsourced. Based on the production cost, they extend this cost measure to OSS as a useful asset used in production.

The paper provides estimates of U.S. annual OSS investment based on the prices prevailing during the period the investment took place (nominal) and adjusted for inflation and quality changes (real), which allows for comparisons across time. In addition to nominal and real investment series, the paper includes estimates of the net-stock series (current cost of replacing all available assets at the end of the year) for the 2009–2019 period. The estimates indicate that the U.S. investment in 2019 was \$37.8 billion, with a current-cost net stock of \$74.3 billion.

The authors conclude that the results provide a strong baseline for resource cost estimates, and their goal is to repeat the process each year to provide a consistent picture of how the OSS ecosystem evolves over time. They have made the tools and datasets they developed publicly available to encourage further research and analyses into the measurement of OSS and its contribution to productivity.

This *Research Spotlight* was prepared by *Survey of Current Business* staff. It uses language from the article “[From GitHub to GDP: a framework for measuring open source software innovation](#)” by J. Bayoán Santiago Calderón, Ledia Guci, Gizem Korkmaz, Brandon L. Kramer, and Carol A. Robbins, published in the April 2024 issue of *Research Policy*. An [earlier version](#) was published on the BEA website as a working paper in July 2022.

The views expressed in this *Research Spotlight* are those of the authors and do not necessarily represent the U.S. Bureau of Economic Analysis or the U.S. Department of Commerce.



Subscribe to the *SCB*

The *Survey of Current Business* is published by the U.S. Bureau of Economic Analysis.
Guidelines for citing BEA information.

Survey of Current Business

bea.gov/scb
scb@bea.gov