

Valuing the U.S. Data Economy Using Machine Learning and Online Job Postings
José Bayoán Santiago Calderón (Bureau of Economic Analysis), Dylan G Rassier (Public
Company Accounting Oversight Board)

With the recent proliferation of data collection and uses in the digital economy, the understanding and statistical treatment of data stocks and flows is of interest among compilers and users of national economic accounts. In this paper, we measure the value of own-account data stocks and flows for the U.S. business sector by summing the production costs of data-related activities implicit in occupations. Our method augments the traditional sum-of-costs methodology for measuring other own-account intellectual property products in national economic accounts by proxying occupation-level time-use factors using a machine learning model and the text of online job advertisements (Blackburn 2021). In our experimental estimates, we find that annual current-dollar investment in own-account data assets for the U.S. business sector grew from \$84 billion in 2002 to \$186 billion in 2021, with an average annual growth rate of 4.2 percent. Cumulative current-dollar investment for the period 2002–2021 was \$2.6 trillion. In addition to the annual current-dollar investment, we present historical-cost net stocks, real growth rates, and effects on value-added by the industrial sector.