Making AI Count: The Next Measurement Frontier

Draft chapter for NBER volume *Economics of Transformative AI: A Research Agenda*.

Diane Coyle & John Lourenze Poquiz, Bennett School of Public Policy, University of Cambridge

Revised September 2025

As generative Artificial Intelligence (AI) advances, it is expected to drive profound transformations in the structure of production and consumption, beyond those already brought about by digital technology, as well as accelerating economic growth. However, capturing AI's contributions within existing measurement frameworks presents significant challenges, both familiar and new, and it will be challenging to see signs of transformation in current statistics.

The 2025 System of National Accounts (SNA25) recognizes the role AI as a transformative technology by including classification it as a distinct type of software. However, there are many challenges in reflecting adequately AI activities and assets in economic statistics, beyond basic data collection, including asset valuation, the construction of price deflators, and depreciation methods. Beyond these well-known issues, broader questions remain about how to measure AI's direct and indirect effects on output, productivity, and welfare, particularly given the platform-mediated, rapid, and potentially transformative diffusion of a general-purpose technology.

This chapter focuses on AI's transformation of economic structures. To trace this, we will discuss measurement along the entire AI 'supply chain': inputs to AI across the stack; AI outputs; the measurement challenges transformative AI creates across the rest of the economy; transformation of business processes; transformation of work and consumption; and ultimately the measurement of productivity and welfare. Many of these issues are not conceptually new (although some are), but they will be significantly exacerbated by transformative AI. Our intention is to identify what economic statistics could reveal as AI becomes transformative.

1. What is the measurement problem?

The adoption of large language models (LLMs) has rapidly become widespread, spanning a broad range of applications, from professional tasks such as drafting emails and generating presentations, to everyday activities like travel or meal planning. As generative AI technologies and applications continue to evolve, their usage is expected to expand significantly. To what extent are official macroeconomic aggregates, such as GDP and productivity statistics, capable of capturing the value provided by AI? The answer is that they will serve poorly: conventional economic statistics tell the story of transformation only with long lags, as classifications, data gathering and even concepts need updating to reflect fully the reshaped economy.

In principle, many AI-related services are recorded in official statistics through observable market transactions, such as subscription fees or usage-based payments via application programming interfaces (APIs). However, the widespread availability of "free" services offered by AI providers complicates the measurement of the value generated for users. For instance, as of May 2025, OpenAI reported approximately 15.5 million subscribers to its standard paid service.[1] This suggests that only a small proportion—around 3%—of its estimated 400 million weekly active users are directly contributing to measured market output, with the remaining 97% accessing free versions. While it can be argued that some of the costs of providing services to free users are subsidized by paying customers, as in the case for most multisided platforms, the standard macroeconomic accounting framework would not explicitly reflect these activity on the side of household consumption, which would limit any analysis seeking to measure the benefits of AI to consumers.

This problem is not new. Similar challenges have long existed in the treatment of digital services provided at zero monetary cost to consumers, such as search engines, social media platforms, and open-source software. While these services generate significant consumer benefits and enable productivity-enhancing activities, they are generally not explicitly accounted for on the consumption side of the national accounts due to the absence of

[1] See NerdyNav Team. "ChatGPT Statistics and Trends (2024)." NerdyNav, 2024. Available at: https://nerdynav.com/chatgpt-statistics and Jessica E. Lessin. "ChatGPT Subscribers Nearly Tripled to 15.5 Million in 2024." The Information, June 6, 2024. Available at: https://www.theinformation.com/articles/chatgpt-subscribers-nearly-tripled-to-15-5-million-in-2024

market transactions with their users. The growing prevalence of generative AI will greatly exacerbate this gap.

Several solutions have been proposed, including recording the cost of service provision as part of final consumption, employing stated preference and binary choice experiments, or using market prices of paid services as proxies for the value of free versions (Coyle 2025). These approaches aim to capture the value received by households from the consumption of digital services. However, one of the most significant promises of AI lies in its potential to enhance workplace efficiency and transform production processes across a wide range of industries. In this case, AI services are consumed by firms as intermediate inputs. This will affect the construction of input-output and supply and use tables. These tables tracing outputs from one set of industries as inputs used by other industries are useful for the analysis of structural changes such as the shift from manufacturing to services (United Nations Department of Economic and Social Affairs 2018). The use of free services will make it harder to trace how the tech sector impacts other industries, its "forward linkages". But beyond this, the use of generative AI will not necessarily be captured in the input-output framework as it will change processes as much as the inputs and outputs. For example, a business might continue to purchase AI-enabled accountancy or legal services in the process of producing its own AI-assisted pharmaceutical products. The input-output relationship will not necessarily show the transformation, but rather, eventually, its consequences. By analogy, the arrival of steam engines enabled the factory system, but input output tables (had they existed at the time) would simply have shown increased volumes over time of both steam engines (as an input) and cotton (as an output) without rendering visible the transformation in the mode of production. For that, a count of factory construction and urban textile workers would be needed.

A first step is to collect administrative or transactions data to identify what AI services firms are purchasing, which statistical agencies are currently not doing. Current classifications are too broad to isolate AI activity, and in existing surveys the purchase of AI services tends to be lumped together with the purchase of other software services. Other indicators of diffusion (such as the UK survey results summarized in Table 1) are either not regular or detailed enough to monitor change. These surveys typically record whether firms adopt AI but not the value or type of AI services used, information essential for measuring AI's role as an intermediate input and for assessing its impact on production. They provide extensive margins but not intensive margins of use.

Table 1: Rates of AI adoption by industry, UK

	AI Not applicable		Not Using AI		Using and Testing AI	
	ONS - MES	DSIT	ONS - MES	DSIT	ONS - MES	DSIT
Agriculture, Mining, Manufacturing, & Utilities	64%	53%	21%	31%	14%	16%
Construction	74%	65%	14%	24%	12%	11%
Distribution, hotels & restaurants	65%	61%	19%	27%	16%	12%
Transport, storage, & communication	46%	35%	20%	31%	34%	34%
Finance, Real Estate, and Business Services	49%	38%	23%	35%	29%	27%
Other services	60%	53%	20%	32%	20%	15%
Total	60%	51%	20%	30%	20%	19%

Note: Table compares the rates of AI adoption by aggregate industry classifications based on the 2023 Management Expectation of the Office for National Statistics and AI adoption survey conducted by the UK's Department for Science, Innovation and Technology in 2024. Shares were calculated by the authors using unweighted counts for comparability.

What’s more, the nature of AI products complicates efforts to account for their services as intermediate inputs. Some are embedded in tools such as search engines (e.g. Google), coding platforms (e.g. GitHub), and collaborative environments (e.g. Overleaf) whose *usage* is largely not reflected as industry inputs in macroeconomic aggregates and software such as Microsoft 365 and Adobe Acrobat. This can be viewed as a form of quality improvement, which should ideally be captured through adjustments to price deflators. However, even this would still fail to address the core issue: official statistics do not adequately reflect the flow of AI services across industries.

While AI’s macroeconomic footprint can be inferred through measured trends in software investment, software R&D, or TFP, treating AI as an intermediate input would improve the visibility of sectoral adoption and enable productivity analysis. In the UK and EU, software R&D is currently aggregated with software copies, limiting the analysis of AI-related investment. At the minimum, disaggregating these components would be a critical step but in any case expanded data collection would be needed to trace transformative effects of AI in production.

These difficulties in accounting for AI services as intermediate inputs are further compounded by the growing capability of generative AI to produce creative outputs and other intellectual property that may themselves be capitalized. Examples include software, design assets, written content, and audiovisual materials, all of which can now be generated rapidly and at scale, often with minimal human intervention. This is already happening to some degree, for example with multinationals like Coca-Cola using AI Art as part of its advertising strategy (Marr 2023). One can easily imagine that this could be extended to AI-generated advertising jingles, or even full movies or entire television series are generated by AI. It is simply unclear how to conceptualise these phenomena in the current national accounts framework.

Traditionally, such creative outputs used in production over several periods are treated as capital formation and valued either at production cost or market price. However, AI-generated content is frequently produced at near-zero marginal cost to the user, undermining the validity of cost-based valuation methods. Market-based approaches are equally problematic, particularly given the ambiguity in defining the unit of output and the highly context-specific nature of valuation. For example, what constitutes the unit of measure for a song? Its duration? Complexity? Should AI-generated songs be valued on par with works by established artists like Taylor Swift or Oasis? What if the melody is the same as any other Oasis song?

Lastly, even AI services that are currently captured through market transactions can pose significant measurement challenges due to rapid quality improvements. Many AI systems improve through continued use, user feedback, and fine-tuning, leading to substantial gains in performance over time without corresponding changes in price. This dynamic complicates the application of traditional price deflators, which often assume stable product characteristics or rely on observable input costs (Coyle 2024). As a result, standard methods for deflating nominal output may fail to capture the true productivity gains associated with AI.

Given the limitations of the conventional statistical framework, we next consider alternative approaches to measuring the impact of transformative AI, using the concept of the AI value chain as an organizing structure. One of the messages is that no single lens will make visible the whole picture of the transformation.

2. Accounting for AI Inputs

Much like the challenges associated with measuring AI services, accounting for the inputs to AI production itself will present significant difficulties. The development and deployment of AI systems rely on a complex mix of inputs that differs from traditional production. This includes large-scale computing infrastructure, specialized software, vast datasets, and ongoing human and organizational resources. These inputs have high upfront costs, are often intangible, distributed across borders, and subject to rapid technological change.

Large-scale investment in ‘hard’ infrastructure such as data centers, chips, network equipment, alongside as software and intangibles such as R&D is essential for enabling the continuing deployment of AI technologies.

One key measurement issue is the geographic dispersion of AI infrastructure. The data centers supporting training and inference may be located across multiple countries, making it difficult to attribute investment and productive capacity to specific national economies. A significant policy challenge is understanding the extent to which national AI systems are exposed to foreign supply chains, especially if they consider AI systems as part of national critical infrastructure. Addressing this issue requires substantial improvements in the granularity and collection of trade statistics, as well as the development of multiregional input-output (MRIO) tables and capital flow accounts that can capture the complex, cross-border nature of AI-related investment and production.

Table 2: Forecast annual growth in power demand by US data centres, to 2030

	Growth projection in power demand
International Energy Agency	7%
Electric Power Research Institute	10%
McKinsey & Co	11%
S&P Global	13%
DC Byte	14%
Barclays	18%

Source: “AI revolution: Meeting massive AI infrastructure demands” by Barclays Investment Bank (2024).

The operation of AI systems also requires a substantial amount of energy. AI models consume significantly more electricity than conventional digital services. Queries often require up to ten times more energy than a standard web search. The infrastructure underpinning AI, particularly large data centers, demands continuous and reliable power supply. Even training AI models require a substantial amount of power. Analysts estimate a 7% to 18% annual increase in US energy demand to 2030 due to data centers (Table 2).

Future innovation in both chip design and AI development will likely aim to economize on energy (and water) use; but meanwhile growth in the readily-measured output of electricity at national level is a useful indicator of AI usage.

Mitu and Mitu (2024) show that even training AI models produces hundreds of tonnes of CO_2 emissions. These environmental costs are often incurred where infrastructure is located, not where services are consumed, yet environmental accounting remains largely national. This misalignment highlights the need for international frameworks that better reflect the global footprint of AI. The same issues apply to water use as well, used for cooling data centers (OECD.AI 2024). The "design of global accounts that incorporate data within and beyond national jurisdictions", is included as part the research agenda of the SEEA Ecosystem Accounting (SEEA EA) framework (2021). As AI continues to scale globally, the need to reconsider how national statistical systems track transboundary energy use and emissions will become more pressing. Existing efforts such as the OECD's inter-country input-output tables,[2] the European Commission's Exiobase,[3] and the EU's FIGARO[4] (Full International and Global Accounts for Research in Input-Output Analysis), offer starting points, but must be expanded and harmonized.

---

[2] See Organisation for Economic Co-operation and Development (OECD). Inter-Country Input-Output (ICIO) Tables Dataset. OECD.Stat. Available at: https://www.oecd.org/en/data/datasets/inter-country-input-output-tables.html

[3] See European Environment Agency (EEA). EXIOBASE: A Global Multiregional Environmentally Extended Supply and Use / Input-Output Database. Available at: https://www.eea.europa.eu/data-and-maps/data/external/exiobase

[4] See Eurostat. Supply, Use and Input-Output Tables (SUIOT) Database. European Commission. Available at: https://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/database

In terms of labor inputs, official data lack the granularity to identify workers developing or operating AI systems. AI roles cut across traditional occupations, combining software skills, data science, and domain expertise, and are frequently embedded in broader functions. It is therefore difficult to isolate the AI workforce using standard occupational classifications. Recent efforts, such as Calvino et al. (2024), attempt to measure industry exposure to AI human capital by analyzing job vacancy data using keyword-based classification. While innovative, these approaches are limited by the availability and representativeness of job postings (Saad et al. 2023; Vassilev, Romanko, and Evans 2021; Ao et al. 2023; Carnevale, Jayasundera, and Repnikov 2014). They also assume that listings accurately reflect the skills required in practice; but, for example, in some instances vacancies listings claim a requirement for multiple programming skills for a role that primarily involves spreadsheet work.

Addressing this gap could involve improving labor force surveys with more detail on tasks, time devoted to tasks, and technologies used. This could be triangulating with time use data, administrative records, and employer surveys. Platforms such as LinkedIn or Lightcast, with rich skill and job data, can also provide insights when combined with traditional sources. There are already examples of this kind of approach – although none in regular statistical production. For instance, Ramraj, Sivakumar, and others (2020) and Liu et al. (2019) developed occupational classifications using LinkedIn data. More recently, LinkedIn[5]

[5] See LinkedIn Economic Graph. AI in the EU: Navigating the European Artificial Intelligence Landscape. 2024. Available at: https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/PDF/AI-in-the-EU-Report.pdf

examined the impact of AI on European workforce. Such efforts can help build a more comprehensive picture of the human capital supporting AI production (and operation) and provide more accurate assessments of skill gaps and labor market dynamics. A future challenge as AI diffuses will be identifying the relevant workers; at present, the number of businesses involved in AI production is relatively small, certainly at the technology frontier. The AI-producing sector will grow over time, and will encompass specialist producers as well as those producing general models. One possibility is to go a step back in the human capital chain and look at the number of PhDs being granted in AI-related fields AI-related fields. Perhaps a more comprehensive approach would be to trace the broader set of activities and institutions involved in AI knowledge production, linking data on grants, publications, patents, and career transitions to map how talent, ideas, and capabilities accumulate and flow across sectors and over time (Lane 2023).

Lastly, AI considerably complicates standard ways of accounting for data inputs. The 2025 SNA recommends that data henceforth be recorded as an asset. As with many intangibles, there are significant measurement challenges. One key question for this exercise is how to value data. Coyle and Manley (2023) provide a review of the many possible approaches. However, the SNA 25 recommends the use the sum of cost for the valuation of data assets (United Nations Statistics Division, n.d.). The cost of data production includes labor, intermediate inputs, capital consumption, and, for market producers, a mark-up reflecting the expected future profitability of the data. Currently, this sum-of-costs approach is regarded as the most feasible method for valuing data assets within official statistics

(Organisation for Economic Co-operation and Development 2015; 2022). It is consistent with how other own-account intangibles, such as R&D, are estimated. But it is far from perfect. Among other drawbacks, as AI applications expand and model performance improves, the value of training data changes accordingly. This introduces a circularity, whereby the value of the input (data) is influenced by the value of the output (AI-driven services or products).

These challenges underscore the need for further methodological research to improve the valuation of data and other intangible assets in the context of national accounting. One possible avenue involves the use of experimental or quasi-experimental methods to assess how variations in training data affect model performance, thus allowing analysts to infer the relative economic value of different datasets. Another prospect is the development of data exchanges. Although in their infancy, countries from China to the UK[6] are experimenting with establishing exchanges, which may evolve to deliver market prices for data as the units of data exchanged are standardized. It is notable that there are few existing data markets but those that exist – such as financial data markets or advertising data markets – are built on standardized units.

Any approach will demand interdisciplinary collaboration among economists, statisticians, computer scientists, and other domain experts, as well as data feeds of granular information

[6] See UK Government. Industrial Strategy: Building a Britain Fit for the Future. Department for Business, Energy & Industrial Strategy (BEIS), 2017 (or latest update). Available at: https://www.gov.uk/government/publications/industrial-strateg

about the nature, source, and structure of the data used in AI development. Since firms often treat training data as proprietary, it could be difficult to require sufficient information to support this type of empirical evaluation. Developing the practical and legal framework and statistical instruments capable of accounting for these complexities will be essential for improving the measurement of data as a production input and for aligning valuation practices with the realities of modern AI systems.

3. AI outputs and broader Impacts

Perhaps one of the most significant measurement challenges posed by transformative AI is capturing its wide-ranging impact across other industries – which is exactly what is needed to track transformation. As a general-purpose technology, AI will influence not only productivity levels but also the quality and nature of outputs in a broad array of economic activities. This is particularly evident in sectors such as healthcare, finance, education, and creative industries, where AI is rapidly improving service quality, diagnostic accuracy, content generation, and personalization. In many cases, these developments alter the structure of economic activities themselves—what constitutes "instruction," "care," or "creative work" may shift—disrupting established statistical classifications and in addition the boundary between market and home production. Tasks will change as new ones are created, and the processes in which they are embedded will change too. These challenges are not all new: the measurement of quality change, service innovation, and intangible outputs has long been a known difficulty in national accounts. However, the rapid and pervasive deployment of AI technologies significantly amplifies their scope and urgency, while some challenges – such as the definition of activities – are new.

One core difficulty lies in measuring changes in product and service quality. Traditional deflators and output measures are typically constructed to capture changes in quantity or price, not in functionality or user experience. In healthcare, for instance, AI-assisted diagnostics and predictive models can improve outcomes and reduce errors, leading to real welfare gains without proportionate changes in observable inputs or expenditures – or indeed with reductions in expenditure. There is already systematic evidence that AI is changing the health sector substantially (Ullah and Ali 2025; Choudhury and Asan 2020; Koo et al. 2024; Ayorinde et al. 2024). However, output in this sector is often measured using cost or revenue data, which may not necessarily reflect improvements in quality. As a result, advances due to AI, such as greater diagnostic accuracy or more effective interventions may go unmeasured.

Similar dynamics are evident in consumer services. AI-powered chatbots, recommendation engines, and virtual assistants enhance convenience, responsiveness, and customization. The nominal value of service output is often measured using firm revenues. In the national accounts, changes in quality should be treated as part of volume change. So, to arrive at estimates of real output the price index needs to be adjusted for changes in quality. National statistical institutes try to account for this using matched models and hedonic approaches. While the practical implementation of these techniques is rarely straightforward, they are more often applied to goods whose product characteristics are observable and relatively stable (such as cars). AI services quality improvements will be intangible and rapidly evolving. Yet quality adjustment would be necessary for measuring AI-driven productivity growth. A new issue with transformative AI will be how to account for the emerging agentic workforce serving consumers and businesses. Expenditure on software or services will give some indication of the growing role for AI but will not capture the fundamental changes in characteristics.

There is virtually no work in this area within official statistics, leaving a growing gap between the lived experience of service enhancements and how they are captured in economic data (Coyle 2025). One practical way to address this is by leveraging big data sources—such as marketing analytics, platform usage metrics, and consumer sentiment—to construct indicators of perceived quality improvement. This implies the need to experiment with novel approaches to developing quality-adjusted price indices, building on prior work, for example using web scraped prices and scanner data (Feenstra and Shapiro 2007; Ivancic, Diewert, and Fox 2011; Białek and Beręsewicz 2021; De Haan and Krsinich 2014).

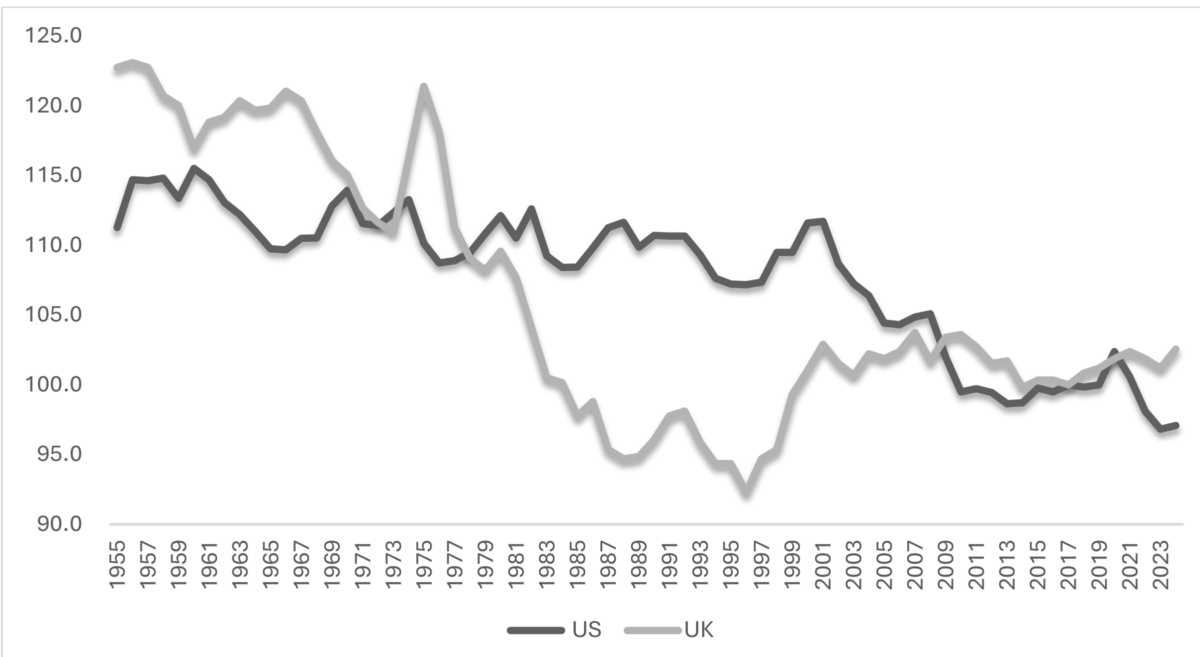
Moreover, while AI can drive significant quality improvements, it also generates new externalities that are not accounted for in existing frameworks. These include negative externalities, such as algorithmic discrimination, which could affect pricing. Varian (2018) explains that the wide spread use of AI provides opportunities to adjust prices based on customer characteristics. Greater personalization will prove problematic for the standard approach. Statistical agencies already struggle to capture price changes when they vary based on individual consumer characteristics rather than broader categories like age or geographic location. For example, if an airline tailors ticket prices according to a user's browsing history or online behavior, determining a consistent "average" price becomes difficult.

In the public sector, where outputs are often in any case valued by the cost of inputs, the diffusion of AI could lead to significant underestimation of value. AI-enabled tools in services such as education, welfare administration, and policing can save time by reducing administrative burdens, enhancing citizen engagement, and providing targeted solutions (Smith 2024; Local Government Association 2025). While there is already work trying to measure quality improvements—particularly in sectors like education and health—the focus has often remained on quantifying outputs rather than evaluating outcomes; yet the benefits of AI in such services stem from enhanced effectiveness rather than increased service volume. Developing metrics that reflect these outcome-based improvements is essential to ensure that national accounts and productivity statistics adequately capture the economic value created by AI-driven innovations.

4. Process changes

The rise of transformative AI will significantly alter how production takes place, reshaping the complementarities between labor and capital. Traditional growth accounting frameworks rely on relatively stable distinctions between capital services, labor input, and total factor productivity, as well as a relatively stable production function. However, as AI systems increasingly substitute or augment human tasks, this conventional construct is breaking down. AI infrastructure, once developed and deployed, can deliver productive services across firms and sectors without a corresponding increase in measured labor input. This makes it difficult to attribute output growth accurately to specific input contributions, and therefore difficult to interpret productivity measurements.

Figure 1. Labor share index (2019 = 100) for the US and the UK



Source: Federal Reserve Economic Data and the UK's Office for National Statistics

From a macroeconomic perspective, this highlights the need to pay close attention to the labor share. The labor share has fallen significantly since the 1950s (tab 1) although more stable in recent decades. Trammell and Korinek (2023) highlight that advanced AI may lead to capital-biased technological change, with gains from automation accruing disproportionately to capital. Using AI patent data, Minniti and co-authors (2025) show that regional AI innovation is strongly associated with declines in regional labor shares. This shift has important implications for income distribution and macroeconomic dynamics, For example, the rollout of driverless taxis and autonomous delivery vehicles replaces human drivers with capital-intensive AI systems, transferring income from labor to firms that own the underlying technology. In media and marketing, generative AI tools are increasingly being used to create content that was once produced by freelance writers, designers, or voice actors. In each case, tasks previously performed by workers are automated, but the value they generate remains within the firm, often without corresponding increases in measured employment or wages, making the labor share something to keep an eye on.

It is important to note that in most of these cases, the tasks themselves remain largely unchanged; it is the factors of production that are shifting. This highlights the need to collect and analyze task-based information to better understand how work is being reorganized across sectors, even when job titles or output metrics appear stable. This includes metrics of the skill-level or expertise involved in tasks, to assess the extent to which AI is a complement or a substitute for them (Autor and Thompson 2025; Restrepo this volume). The descriptors of tasks will need to be updated as the AI transformation progresses.

However, the way in which tasks are organized will also change. Consider, for example, a news organization where the traditional production function involved numerous reporters and writers generating content. With the introduction of generative AI, much of the writing can now be automated (Thomson Reuters Foundation 2025), leaving a smaller number of editors to curate, refine, or fact-check the output, in an altered workflow, and with the introduction of the new factor of data. Similarly, in higher education, tasks such as exam marking could soon be handled by AI tools (Gobrecht et al. 2024; Codiste Pvt. Ltd. 2024), with human oversight limited to final review. In both cases, the same task is still being performed, but by a different mix of inputs combined in a different organizational ‘technology’. Job and task classifications will fail to capture such reconfigurations, when content and process of work has been fundamentally altered.

For example, a news organization may choose to hire fewer reporters, resulting in a recorded decline in individuals classified under occupation code 249202 (Journalists and Newspaper Editors. However, many of the tasks once carried out by junior writers, such as drafting routine articles, are now being performed by AI systems, with a smaller team of editors overseeing, curating, and refining the output. While the occupational title remains the same, the nature of the job and production process in which individuals work is changing significantly. There is a shift in skill requirements, from content generation to editorial judgement, fact-checking, and prompt engineering. To address this, improvements in time-use surveys and job-task mapping are needed to track how AI is redistributing effort within occupations and across sectors. These shifts could happen fast, certainly faster than statistical agencies are able to adjust. New data sources will need to be sought, while time use at work seems a useful dataset to collect.

Transformative AI is reshaping business models by shifting efficiency gains from labor and capital to algorithmic decision-making and automation. These gains often come from internal process improvements such as faster turnaround, better inventory management, reduced downtime, or fewer errors, rather than visible increases in output or revenue, or in an increase in either labor or physical capital inputs.

For example, major firms across sectors are already realizing AI-driven efficiency gains through internal process changes (De Silva 2025). UPS has deployed LLMs to automate customer service interactions and is exploring robotics to streamline parcel handling. Walmart and Target use AI systems to forecast demand and optimize inventory, reducing stockouts and improving replenishment accuracy without increasing headcount. Qantas has implemented AI to optimize flight routes, achieving significant fuel savings through better scheduling and resource use (The Australian 2023). In standard macroeconomic accounting, these improvements are typically captured as changes in Total Factor Productivity (TFP), which is measured as a residual after accounting for labor and capital inputs. Because TFP is a residual, it provides little insight into the underlying drivers of productivity gains, and would also lag the changes occurring in the economy. It is important to identify and quantify the specific mechanisms through which AI enhances efficiency.

This is all the more so transformative AI may deepen the paradox of efficiency: measured output could fall even as welfare rises. By automating tasks like scheduling, customer support, and admin processes, AI eliminates inefficiencies that previously contributed to measured GDP via labor or transactions. As these vanish, productivity improves but measured output may decline.

Yet national accounts measure production, not welfare. For example, if AI enables accurate medical triage via chatbots or speeds up legal services, recorded output may fall despite equal or better outcomes. These efficiency gains would reduce recorded transactions, potentially slowing down GDP and productivity. Ultimately, improvements in consumer surplus, time savings, and service accessibility may go unmeasured or be misclassified as economic decline. Hulten and Nakamura (2017) describe this as output-saving technical change. Given that current official statistics have failed to capture the digital transformation of the economy to date, the AI transformation will exacerbate their shortcomings. New approaches to measurement will be needed. While general equilibrium effects may eventually raise total output, the short-run impact may be a decoupling of welfare improvements and measured output, reflecting a known limitation of GDP as a welfare indicator rather than a paradox of productivity itself.

5. Time

One new approach is capturing the time savings enabled by AI. The process improvements that have historically driven dramatic productivity gains have often involved speeding up production, from sailing by steam instead of wind to the just-in-time automation revolution in manufacturing.

AI's automation capabilities have the potential to reshape not only the processes of production but also the allocation and experience of time across work, leisure, and consumption. By reducing the need for human intervention in routine or repetitive tasks, AI systems can significantly lower the time input required for both market and non-market activities. These time savings, while often difficult to quantify, represent a core channel through which AI contributes to productivity growth and economic welfare. Likewise, the welfare benefits of AI may be most visible in how it alters people's use of time, especially in terms of reduced administrative burden, faster service delivery, and more efficient decision-making

Chang (2010) argued that the washing machine changed the world more than the internet, directly altering how people, especially women, spent their time. The broader point is that technological progress should be evaluated not only by its market output, but by how it transforms everyday life and reconfigures human activity. AI may follow a similar trajectory. AI promises large time savings in cognitive, administrative, and logistical tasks—both in professional settings (e.g. automating emails, scheduling, customer queries) and personal life (e.g. travel planning, content filtering, personal finance). This creates new possibilities for reallocating time toward higher-value tasks, leisure, or care work, with substantial implications for labor markets, gender dynamics, and well-being. AI is now doing the same for white-collar and knowledge-intensive sectors.

One of the more profound implications of AI lies in its potential to free people from routine cognitive tasks, creating space for higher-order thinking, creativity, and innovation. By automating repetitive workflows such as data cleaning, literature searches, or coding routines, AI can shift the focus of knowledge workers toward more conceptual and strategic activities. For instance, an economist who previously spent considerable time writing and debugging code can now use AI tools to automate much of that work, allowing more time to refine conceptual models, explore alternative specifications, or develop new theoretical insights (Korinek, J Ec Literature survey; Mullainathan, this volume). This reallocation of cognitive effort echoes historical shifts in manual labor, where mechanization freed workers to engage in more skilled or supervisory roles.

How could metrics capture the shift from routine to higher-order cognitive work? Traditional labor metrics, such as hours worked or employment counts, may remain unchanged even as the nature of tasks evolves significantly. Time use surveys offer one avenue for incorporating these effects, but they are often infrequent, coarse-grained, and disconnected from production statistics. A richer integration of time-use data into economic measurement—particularly with respect to digital service provision, remote work, and task automation—could help to better reflect the welfare gains generated by AI.

On the household side, AI is already enhancing domestic efficiency in mundane tasks, and future advancements suggest even greater transformation. Today's AI-powered vacuum cleaners can not only clean but also intelligently identify and pick up small items, adapting to changing environments (Business Insider 2024). Within households, emerging humanoid robots are being deployed in trial homes to assist with chores and provide companionship, marking a shift toward more generalist robotic helpers (Johnson 2024; Vincent 2024). Other efforts are pushing robots toward more sophisticated, whole-body manipulation capable of tasks such as folding laundry or organizing rooms (Jiang et al. 2025). Looking ahead, these household robots promise to take on a broader variety of home production activities freeing up significant time for occupants.

To better capture the welfare gains from these advances, it will be important to develop dedicated measures of household productivity and household capital. As AI and robotics begin to substitute for human effort in home production, traditional approaches to valuing unpaid work such as the replacement cost method, which uses equivalent market wages may no longer be sufficient. One of the key features of current household satellite accounts is their reliance on labor-based valuation, implicitly assuming that household output is primarily driven by human input. However, as returns to capital would begin to account for a larger share of value added in this area. This mirrors the challenge discussed earlier in valuing own-account data in section 2.

Accounting for the full extent of time-related welfare gains enabled by AI, traditional statistical sources will need to be supplemented with alternative data. New forms of digital trace data such as geolocation, activity logs, and interaction timestamps offer valuable potential for understanding how time is reallocated across tasks, sectors, and social groups.

For instance, anonymized data from platforms like Google Maps could potentially help measure changes in commuting time, reflecting shifts in work patterns enabled by remote work technologies and AI scheduling tools. Similarly, metadata from app usage, wearable devices, or smart assistants could offer novel indicators of task duration, intensity, and frequency.

Social media content and other user-generated data sources can also provide indirect evidence of changing time use and consumption preferences. While not without methodological challenges (including biases in representation) this type of data can shed light on how individuals experience time savings or respond to AI-enabled service enhancements. For example, scraping content related to healthcare appointments, customer service experiences, or online learning platforms could provide qualitative and quantitative insights into where AI is making processes faster or more accessible.

Many of the most relevant data are held by private technology firms whose AI products generate large-scale behavioral data across users and contexts. Carefully governed data-sharing partnerships between tech firms and national statistical agencies could provide a mutually beneficial model for expanding measurement capabilities, enabling a better public understanding of the economic potential of transformative AI. Such collaborations could offer insights into aggregate time patterns without compromising user privacy, especially if standardized APIs and anonymization protocols are used. These efforts would allow for the development of new indicators that track time efficiency, service responsiveness, and other dimensions of AI-enabled productivity, thus helping to bridge the gap between welfare gains and official economic statistics. As AI continues to reshape the temporal organization of the economy, incorporating these alternative data sources will be essential for a more complete and timely understanding of its impact.

6. Summary and final remarks

This chapter has outlined some of the key challenges that transformative AI poses for economic measurement. In table 3 we summarize these, distinguishing existing (although exacerbated) measurement challenges and novel ones. As AI reshapes both production and household activities, it complicates practically all aspects of economic measurement within the existing framework, as well as posing new conceptual challenges. In order to capture these changes, more granular, task-based, and outcome-oriented measures will be essential to ensure that statistics remain relevant and informative in an AI-driven economy.

Table 3: Measurement challenges posed by transformative AI

		Measurement challenge	Possible Indicators
AI inputs	Existing	Cross-border energy use	Extended multi-regional IO tables (MRIO), CO_2 emission trade account
		Labor input – tasks, skills	Task-based questions in labor force surveys, time use data
	Novel	Data value - feedback loop	Standardized prices for data exchange, firm-level data valuation modules for business surveys
AI services	Existing	Service quality change	User-generated content, performance benchmarks, customer ratings for quality adjustment
		Personalisation	Micro-transaction data, individualized price logs
	Novel	Structure of economic activities	Time use data on task and technology use, AI-specific business registers
		Agentic workforce, service value	
Process changes	Existing	Input mix, occupation changes	Administrative microdata on occupations, matched employer-employee datasets, time use data
		Household activities	Smart device logs, app usage data, user-generated data analytics
	Novel	Process re-engineering, time to produce	Firm-level process metrics, workflow audits
		Relocation of production boundary	Sector reclassification patterns, market vs household task shifts, time-per-output metrics

Official statistical agencies will need to strengthen existing data collection methods, but will also inevitably need to incorporate new tools and data sources. Transformative AI will disrupt statistical production, as it will so many other activities.

For this to be feasible, robust policy frameworks are needed to facilitate secure data sharing while confidentiality and commercial sensitivities. Trust will be needed among stakeholders so they have confidence data will be used for the public good, respecting confidentiality requirements and privacy. Without such trust, in addition to substantial internal re-engineering, the legitimacy of statistical systems risks being undermined at precisely the time when timely and accurate measurement is most needed to make visible the transformation being brought about by AI.

References

Ayorinde, A, D O Mensah, J Walsh, I Ghosh, S A Ibrahim, J Hogg, and F Griffiths. 2024. "Health Care Professionals' Experience of Using AI: Systematic Review With Narrative Synthesis." *Journal of Medical Internet Research* 26:e55766. https://doi.org/10.2196/55766.

Calvino, Flavio, Hélène Dernis, Lea Samek, and Antonio Ughi. 2024. "A Sectoral Taxonomy of AI Intensity." https://www.oecd.org/.

Choudhury, A, and O Asan. 2020. "Role of Artificial Intelligence in Patient Safety Outcomes: Systematic Literature Review." *JMIR Medical Informatics* 8 (7): e18599. https://doi.org/10.2196/18599.

Codiste. 2024. "AI in Grading: Transforming Education with Intelligent Assessment." https://www.codiste.com/ai-in-grading.

Coyle, Diane, and Annabel Manley. 2024. "What Is the Value of Data? A Review of Empirical Methods." *Journal of Economic Surveys* 38 (4): 1317–37.

Coyle, Diane. 2025. *The Measure of Progress*. Princeton University Press.

Coyle, Diane. 2024. Productivity Measurement: New Goods, Variety and Quality Change. Ch.4 in The Measure of Economies: Measuring Productivity in an Age of Technological Change. Eds. Marshall B Reinsdorf & Louise Sheiner. University of Chicago Press.

Gobrecht, Alexandra, Felix Tuma, Moritz Möller, Thomas Zöller, Mark Zakhvatkin, Alexandra Wuttig, Holger Sommerfeldt, and Sven Schütt. 2024. "Beyond Human Subjectivity and Error: A Novel AI Grading System." https://arxiv.org/abs/2405.04323.

Hulten, Charles, and Leonard Nakamura. 2017. "Accounting for Growth in the Age of the Internet: The Importance of Output-Saving Technical Change." https://www.nber.org/papers/w23315.

Kaldor, Nicholas. 1957. "A Model of Economic Growth." *The Economic Journal* 67 (268): 591–624. https://doi.org/10.2307/2227704.

Koo, T H, A D Zakaria, J K Ng, and X B Leong. 2024. "Systematic Review of the Application of Artificial Intelligence in Healthcare and Nursing Care." *The Malaysian Journal of Medical Sciences* 31 (5): 135.

Lane, Julia. 2023. The Industry of Ideas. AEI.

Local Government Association. 2025. "State of the Sector - AI (2025 Update)." https://www.local.gov.uk/sites/default/files/documents/State%20of%20the%20sector%20-%20AI%20(2025%20update).pdf.

Marr, Bernard. 2023. "The Amazing Ways Coca-Cola Uses Generative AI In Art And Advertising." *Forbes*, September. https://www.forbes.com/sites/bernardmarr/2023/09/08/the-amazing-ways-coca-cola-uses-generative-ai-in-art-and-advertising/.

Mitu, Narcis Eduard, and George Teodor Mitu. 2024. "The Hidden Cost of AI: Carbon Footprint and Mitigation Strategies." https://ssrn.com/abstract=5036344.

OECD.AI. 2024. "How Much Water Does AI Consume?" https://oecd.ai/en/wonk/how-much-water-does-ai-consume.

Smith, Laura. 2024. "How the UK Government's AI Action Plan Encourages Higher Education to Embrace AI." *NIICO.Ai*, March. https://niico.ai/how-the-uk-governments-ai-action-plan-encourages-higher-education-to-embrace-ai/.

Thomson Reuters Foundation. 2025. "TRF Insights: Journalism in the AI Era." https://www.trust.org/wp-content/uploads/2025/01/TRF-Insights-Journalism-in-the-AI-Era.pdf.

Trammell, P, and A Korinek. 2023. "Economic Growth under Transformative AI." https://www.nber.org/papers/w31815.

Ullah, W, and Q Ali. 2025. "Role of Artificial Intelligence in Healthcare Settings: A Systematic Review." *Journal of Medical Artificial Intelligence* 8:24.

United Nations. Department of Economic, and Statistics Division Social Affairs. 2018. "Handbook on Supply and Use Tables and Input-Output Tables with Extensions and Applications." https://unstats.un.org/unsd/nationalaccount/docs/SUT_IOT_HB_Final_Cover.pdf.

United Nations Statistics Division. n.d. "Update of the System of National Accounts (SNA)." https://unstats.un.org/unsd/nationalaccount/snaupdate/dztt.asp.

Varian, Hal R. 2018. *Artificial Intelligence, Economics, and Industrial Organization*. Vol. 24839. National Bureau of Economic Research Cambridge, MA, USA: