AI Watch: Methods and Data

Definitions, sources and limits of each indicator

October 31, 2026 · Gokhan Mergen

This note documents how each number in *AI Watch* is computed. Every indicator is rebuilt from public data by the scripts in `code/`; nothing is read off another publication's chart. Figures attributed to the *Financial Times* or a16z are quoted, and are marked as such in the essay.

S1. Sources, release timing and revisions

Source	Series	Frequency	Typical lag	Folder in `data/`
Indeed Hiring Lab, job postings tracker	Postings index for 47 US occupational categories and for all postings	Daily, updated about weekly	About one week	`Indeed Hiring Lab - job postings by occupation`
BEA, GDP by industry, via FRED	Real and nominal value added: Information; Finance and insurance; Professional, scientific, and technical services; Private industries	Quarterly	About three months	`FRED - output and employment by industry`
BLS, Current Employment Statistics, via FRED	Payroll jobs in the same sectors, total private, and computer systems design	Monthly	About one week after the month	`FRED - output and employment by industry`

Source	Series	Frequency	Typical lag	Folder in `data/`
BLS, Current Population Survey and Productivity and Costs; BEA personal income; all via FRED	Employment rates (ages 25–54 and 20–24), unemployment (20–24), nonfarm business labor share and output per hour, production workers' hourly pay, PCE prices, real disposable income per person	Monthly; quarterly for Productivity and Costs	One week to five weeks	`FRED - labor market, pay and income`
BLS, Current Employment Statistics, via the BLS public API	Payroll jobs in legal services, software publishers and telephone call centers	Monthly	About five weeks	`BLS - employment in AI-exposed industries`
Indeed Hiring Lab, job postings tracker, other countries	Software development postings in Britain, Canada, Australia, Germany and France	Daily, updated about weekly	About one week	`Indeed Hiring Lab - software postings, other countries`
Indeed Hiring Lab, AI tracker	Share of postings mentioning AI, nine countries	Daily, refreshed monthly	About one month	`Indeed Hiring Lab - AI tracker`
Other studies	Figures that cannot be computed here (see S7)	As published		`Quoted figures - hand-entered`

All three sources revise their history. Indeed re-estimates seasonal factors as new data arrive, BLS revises payrolls for two months and benchmarks them each year, and BEA revises value added with each annual update. Each edition therefore uses the vintage downloaded on its date. `code/download.py` writes a `MANIFEST.json` in each data folder recording every file's URL, retrieval time, SHA-256 hash and first and last observation, and `data/SHA256SUMS` lists the hashes. Committing the data with each edition keeps the vintages in Git history.

Indeed's data are published under the Creative Commons Attribution 4.0 licence and must be credited to Indeed Hiring Lab. BEA and BLS data are in the public domain.

S2. Demand for workers: job postings

Measure. Indeed's index of seasonally adjusted job postings, a seven-day trailing average, with February 1, 2020 = 100. "Total postings" (all active postings) is used throughout, not "new postings." Indeed assigns postings to its 47 occupational categories from normalized job titles.

Groups. Indeed publishes indexes, not posting counts, so a group of categories cannot be weighted by its postings. The groups are unweighted means of their categories' indexes:

- Engineering: industrial, mechanical, civil and electrical engineering (Indeed has no chemical engineering category in the current files).
- Healthcare: physicians and surgeons, nursing, personal care and home health, dental, medical technician, medical information, therapy, and pharmacy.
- Production and manufacturing, human resources and software development are single categories.

Because every category equals 100 on February 1, 2020, averaging the indexes weights each category equally at that date. a16z's chart of change since October 1, 2025 averaged the categories' percentage changes instead, which weights them equally at that later date. For engineering the two give +27.3% and +29.5% over October 1, 2025 to October 23, 2026; the second matches a16z's figure exactly, as do its software development (+25.7%) and production and manufacturing (+13.3%) figures, which are single categories. a16z's levels against February 2020 for engineering (about 145) and healthcare (about 131) are lower than the unweighted means (152 and 138) and could not be reproduced; they presumably use other weights.

Changes. The past-year change compares the latest day with the same calendar date a year earlier. Each series' low is its minimum since January 1, 2024, and its peak its maximum in 2022.

Tax timing and other countries. US tax law changed the cost of software development twice in the period covered. The Tax Cuts and Jobs Act (signed December 22, 2017) required research and experimental expenditures, which section 174 defines to include software development, to be capitalized and amortized over five years (fifteen for research abroad) in tax years beginning after December 31, 2021. The One Big Beautiful Bill Act (signed July 4, 2025) restored immediate deduction of domestic research expenditures paid or incurred after December 31, 2024 (new section 174A), with a retroactive election back to 2022 for small businesses; research abroad is still amortized over fifteen years. `postings.py` compares software development postings in the US with Britain, Canada, Australia, Germany and France, where these rules do not apply to local employers, over calendar 2022 (January 3 to December 30) and since July 4, 2025. The

comparison is descriptive: a US-only movement points to a US cause but cannot separate the tax change from other American factors, such as the concentration of AI investment in US firms or the Federal Reserve's policy; and Canada is not a clean control, since US firms' research there faced the fifteen-year rule.

Limits. Postings record intentions to hire, not hires; one posting can be filled by several people or by none; and Indeed's share of all postings varies by occupation and over time. Postings for a category can rise because employers search longer for the same hires. Category assignment rests on job titles, so a job whose content changes but whose title does not stays in its category.

S3. Output and jobs by sector

Sectors. Following the FT's charts of October 30, 2026:

Name in the essay	Industries (NAICS 2022)	Real value added	Nominal value added	Payroll jobs
Tech	Information (51)	RVAI	VAI	USINFO
Finance & professional services	Finance and insurance (52); Professional, scientific, and technical services (54)	RVAFI, RVAPST	VAFI, VAPST	CES5552000001, CES6054000001
AI-adopting sectors	All three			
Rest of the private economy	Private industries less the three	RVAPI	VAPI	USPRIV

Codes are FRED series identifiers. Computer systems design (NAICS 5415, CES6054150001) is reported separately as the closest published payroll series to software services.

Aggregation. Real value added in chained dollars does not add across industries, so each group is a Törnqvist index. With real value added $Y_{i,t}$ and nominal shares $s_{i,t}$ within the group, the group's growth from one quarter to the next is

$$\ln \frac{Y_t}{Y_{t-1}} = \sum_i \tfrac{1}{2} \left(s_{i,t} + s_{i,t-1} \right) \ln \frac{Y_{i,t}}{Y_{i,t-1}}$$

The rest of the private economy is backed out of the private-industries total the same way: its growth is the total's growth less the AI-adopting sectors' share-weighted growth, divided by the rest's average share. Payroll jobs are averaged over the three months of each quarter, and only

complete quarters are used. Both are indexed to 2022 Q4 = 100, the quarter in which ChatGPT launched.

Limits. Payroll jobs count jobs, not people or hours, so a shift toward part-time or longer hours is missed. The value-added series cover private industries including farms, while the payroll series exclude farm workers; the difference is small for these comparisons. Real value added in the Information sector rests on price indexes for software, cloud, telecommunications and media services that are hard to construct, and the sector mixes software and data services with broadcasting and publishing.

S4. Output per worker against trend

Measure. Real value added divided by payroll jobs, indexed to 2012 Q1 = 100, the FT's base. The trend is a straight line fitted by least squares to the logarithm of output per worker over 2012 Q1–2019 Q4, extended to the latest quarter. The gap is the latest value's percentage difference from the trend.

Sensitivity. For the AI-adopting sectors combined, the gap in 2026 Q2 depends on the trend window:

Trend fitted to	Gap in 2026 Q2
2005 Q1–2019 Q4	+15.8%
2010 Q1–2019 Q4	+20.9%
2012 Q1–2019 Q4 (used in the essay)	+21.3%
2015 Q1–2019 Q4	+23.0%

The fitted trend grows 6.7% a year for tech, 0.2% a year for finance and professional services, and 0.5% a year for the rest of the private economy. A nearly flat trend makes the gap for finance and professional services large even for moderate growth, so the essay reports growth rates alongside the gaps.

Limits. A trend fitted to eight years is a reference line, not a forecast of what would have happened without AI or without the pandemic. Output per worker rose across the economy in 2020 because low-paid jobs disappeared first (a composition effect), and some of that persists. These are descriptive comparisons; none of them identifies a cause.

S5. Jobs, pay and the labor share

`code/signposts.py` computes Section 4 of the essay and the young-worker figures in Section 5.

Employment. The employment rate of 25–54-year-olds (LNS12300060): the latest month and the 2025 average, which covers 11 months because October 2025 was not collected. For 20–24-year-olds, participation times one minus the unemployment rate, over period averages.

Labor share. BLS publishes the nonfarm business labor share as an index (2017 = 100). It is shown in percent by scaling the index to the 52.8% level for 2026 Q2 recorded in `data/Quoted figures - hand-entered/quoted.csv`; this assumes that level is consistent with the current vintage of the index. Output per hour is compared over the latest four quarters.

Pay and income. Production workers' real hourly pay is average hourly earnings (AHETPI) deflated by the PCE price index. Twelve-month changes compare the latest month with the same month a year earlier. Real disposable income per person is A229RX0.

The most exposed industries. Payroll jobs in legal services, software publishers and telephone call centers since November 2022, from the BLS API. Output for these industries is not tracked: it would need BEA's detailed industry accounts.

AI in job ads. The share of US postings that mention AI-related terms (Indeed Hiring Lab AI tracker), a broader measure than postings with AI in the job title.

Young workers in exposed jobs. Quoted: the ADP-based estimate of Brynjolfsson, Chandar and Chen, and a Current Population Survey replication (ages 22–25, the two most exposed quintiles against the three least, relative to October–December 2022) computed in the companion repository `cheaper-intelligence-public-data`, with wide intervals.

S6. The data-through line, the claims check and editions

The line under the byline giving the date of the latest data is written by `code/scoreboard.py` between the markers `<!-- latest:begin -->` and `<!-- latest:end -->` in `essay/ai-watch.md`; the rest of the essay is written by hand. `code/claims.py` holds each computed number quoted in the essay, with the sentence around it, and fails if a number no longer matches the results, if the sentence can no longer be found, or if its wording (a month, "fell", "record low") no longer fits the data. After a data update it lists exactly which sentences need revising. `code/update.sh <date>` also saves that edition's headline readings in `results/editions/<date>.json`.

S7. Quoted, not computed

From the FT (John Burn-Murdoch, "Is this the first clear evidence of AI's impact on the labour market?", October 30, 2026): the comparisons for countries other than the United States, the company revenue and headcount series, the junior and senior employment series, and its accounts of the analyses by Eric Basmajian and Goldman Sachs. From a16z (Moses Sternstein,

"Charts of the Week: Job FUD Fades", October 30, 2026): the levels against February 2020 for its engineering and healthcare groups. Copies of both, as supplied by the author, are in `archive/`. Figures quoted from other studies are listed, with sources, in `data/Quoted figures - hand-entered/quoted.csv`.

S8. Scripts

Script	What it does	Writes
`download.py`	Fetches every source listed in `common.py` and records its manifest	`data/`
`postings.py`	Indicator 1: postings by group and category, breadth, a16z check, software postings around the 2022 and 2025 tax changes in six countries	`results/postings.json`
`sectors.py`	Indicators 2 and 3: output, jobs and output per worker by sector, trend, sensitivity	`results/sectors.json`, `results/sectors_quarterly.csv`
`signposts.py`	Section 4: employment, the labor share, pay and income, jobs in the most exposed industries, AI in job ads	`results/signposts.json`
`figures.py`	The five charts, in the house style of `style.py`	`images/*.png`, `images/data/*.csv`
`scoreboard.py`	The data-through line and edition snapshots	`essay/ai-watch.md` (that line only), `results/scoreboard.json`, `results/editions/`
`claims.py`	Checks every computed number in the essay's text	nothing (exit status)
`run_all.sh`, `update.sh`	Recompute everything; download, recompute and rebuild the essay	