

Is AI language showing up in PH job ads yet?

Ralph Torres · Jul 16, 2026 · mail@ralphptorr.es · ralphptorres.github.io/jobshift

2% → 5%

explicit AI terms
($p = 0.0495$)

9% → 10%

AI-or-automation
(not significant)

595

postings scraped
(317 archived, 278 current)

4.5M → 6.9M

matched-sector PH
employment, 2018-25

Method: 595 job postings scraped from Kalibrr, a Philippine job board, across thirty-one general sectors (marketing, retail, logistics, HR, healthcare, education, real estate, IT, finance, e-commerce, media, and others). A pre-ChatGPT snapshot (Oct 2021 to Sep 2022, via the Wayback Machine) is compared against a current one (Jul 2026), checking whether “AI-adjacent” language shows up more in job ads now than before ChatGPT, cross-referenced against PSA’s own employment data.

Finding: a small, real shift.

- Postings mentioning explicit AI terms (AI, ChatGPT, AI tools) rose from 2% to 5% (Fisher’s exact test, $p = 0.0495$).
- A broader, fuzzier “AI-or-automation” bucket rose from 9% to 10%, not statistically significant ($p = 0.89$), consistent with that bucket being diluted by pre-ChatGPT martech and RPA terminology.
- Archived-era AI mentions cluster in already-established fields (data science, RPA). Current-era mentions include new generative-AI tool references (ChatGPT, Claude, MidJourney) absent from the earlier sample entirely.
- PSA’s Labor Force Survey shows total Philippine employment climbing every year since 2020’s COVID dip, ChatGPT’s launch included, so nothing here suggests visible economy-wide displacement even as job-ad language shifts.

The sample spans a general cross-section of sectors rather than one specific industry, so this is a check on whether AI language shows up in job ads broadly, not a claim about any single sector. Whether the shift is larger in a more exposed sector is a separate question, needing a different sampling design.

Method, and two corrections along the way

Kalibrr embeds full listing data (titles, descriptions, qualifications) as structured JSON in the page, so no browser automation was needed. The first pass used only search-result pages (15 jobs each), which the Wayback Machine has very few pre-2023 snapshots of, capping the archived side at roughly 20-30 postings regardless of category count.

Correction 1: switching to individual job-posting permalinks

Wayback holds over a thousand permalinks from late 2021 with the same embedded data, since permalinks get crawled far more often than search pages. That pushed the archived side to several hundred postings, but introduced a new problem: Kalibrr also serves Indonesia (kalibrr.id), and permalink pages are not geo-scoped the way category pages are.

- Roughly half the pull was Indonesian companies (PT Bank Mandiri, Tokopedia) unrelated to the Philippine labor market.
- Filtering to Philippine-resolved locations dropped about 280 Indonesian jobs, leaving 290 clean ones.
- Combined with the original search-page pull: 317 archived postings total. The current-side search was broadened to match, landing at 278.

Correction 2: excluding staffing-agency spam

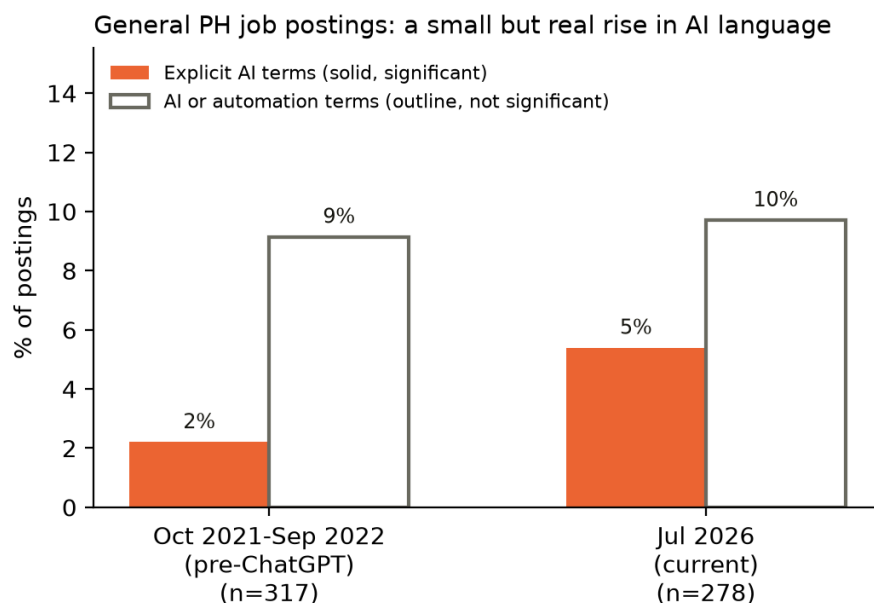
An independent review pass, run to stress-test the method before treating any of it as reliable, found most of an earlier, smaller archived sample was BPO/call-center staffing agencies (Aces Call Center Jobs Inc., Transcom Worldwide, CLEAR Corporate Solutions) cross-posting identical call-center ads under unrelated category tabs like “marketing” and “hospital,” plus one individual posting explicitly titled a BPO role under “management consulting.” All of it was excluded by company name. No single company is more than 9% of either final sample.

Tagging and keyword bugs

Each posting was tagged against two term lists:

- **Explicit AI terms:** ai, artificial intelligence, chatgpt, genai, machine learning, copilot, prompting, prompt engineering, ai tools, ai-assisted, ai-powered.
- **Automation bucket (broader):** automation, automated, rpa. Split out after this bucket kept catching marketing automation (Pardot, Iterable), a 2010s martech category unrelated to ChatGPT-era tooling.

The keyword matching had three regex bugs caught by review. `prompt` alone matched inside `promptly` and ordinary uses like `prompt service`. `rpa` matched inside `counterparts` and `underpaid`. All three are fixed now, either with word boundaries or by requiring more specific phrasing (`prompt engineering` rather than bare `prompt`). None of these are exotic problems. They surface only when someone checks company names and regex matches against source text instead of trusting that the resulting numbers look plausible.



What the postings actually said

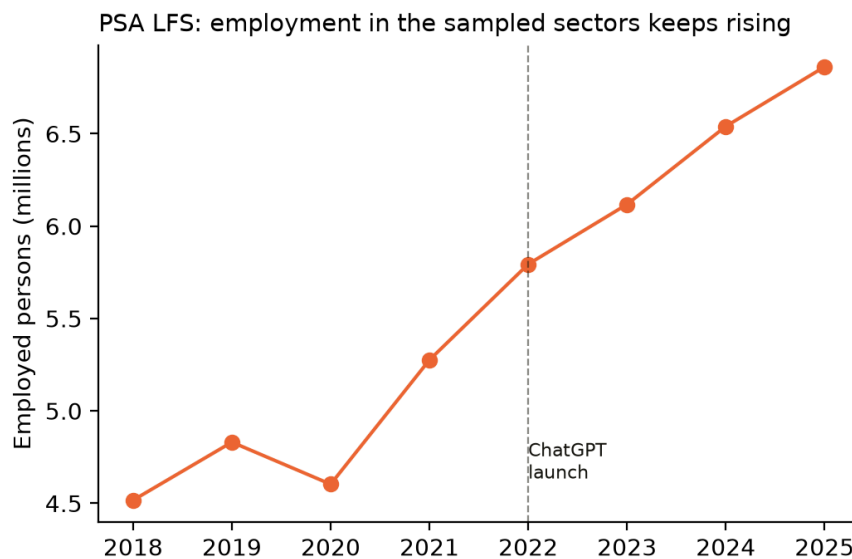
Archived-side AI mentions concentrate in roles already AI-adjacent before ChatGPT: Data Scientist and RPA Developer postings referencing `machine learning` and `artificial intelligence` as domain knowledge, not new tooling. Current-side mentions include the same kind of role, plus new ones impossible in 2021:

- A Junior Multimedia Designer asked to use “MidJourney, Runway, Adobe AI features, ChatGPT for ideation.”
- A Social Media Content Strategist asked for familiarity with “Runway, VEO or Flow, ChatGPT, or Claude.”

- A Machine Learning Specialist role with “{LLM}” in the title itself.

That is the texture behind the 2% to 5% move: not AI replacing the job, but generative-AI tools appearing by name in what the job now expects an applicant to already know. Every posting behind this statistic, with the matched term and surrounding text, is browsable on the site.

A smaller side check: a real 2021 content-writer ad was rewritten “as if posted today” using Claude Sonnet 5, grounded in the actual phrasing of real 2026 AI-mentioning postings from this dataset rather than free-form guessing, with each added line citing the real posting it drew from. The grounded rewrite stays close to the original, adding AI tools as one bonus qualification among several rather than a role overhaul, matching the pattern across the dataset: AI mentions tend to sit inside otherwise unchanged job descriptions. The closest real 2026 match for this specific role (marketing ops, since no content-writer posting turned up at all) still does not mention AI outside company boilerplate.



The bigger picture

Total economy-wide employment is too blunt a comparison for a scrape of specific job categories, so PSA’s Labor Force Survey (pulled directly from PSA’s own PXWeb statistical API, openstat.psa.gov.ph) was narrowed to the six PSIC industry groups closest to the sampled sectors, together covering about 91% of the current-side sample by job function:

- Information and communication (IT, software, media)
- Professional, scientific and technical (marketing, consulting)
- Financial and insurance
- Administrative and support (HR, admin, customer service)
- Education
- Human health

Employment across these six groups rose from 4.5M in 2018 to 6.9M in 2025, recovering from the 2020 COVID dip and climbing every year since, including the two-plus years since ChatGPT launched. This is still an imperfect proxy, since PSIC groups do not map one-to-one onto Kalibrr’s category tags, but it is consistent with no visible mass displacement in the sectors sampled. That lines up with the IMF framing TAF-AIR’s own project brief cites: AI exposure currently looks more like a leading indicator of task-level change than aggregate job loss. Job-ad language shifting while sector employment keeps climbing is consistent with that framing.

Limitations

- This grew well past a few hours of work, but it is still not a study.
- The statistically significant result (explicit AI terms) is real but small, and close enough to the $p = 0.05$ threshold that it should not be overstated. The larger “automation” number is genuinely noisy and needs heavy caveating if cited alone.
- Thirty-one sectors and one job board is not the Philippine labor market, and the archived side depends on whichever categories and postings Wayback happened to crawl in late 2021, not a controlled sample.
- The archived and current sides were also collected by different methods in the end (a permalink sweep plus search pages, versus search pages alone). That asymmetry likely does not bias the direction of the finding, since the two methods mainly affect which postings are visible, not what is written in a given posting, but it is worth naming rather than glossing over.
- A more rigorous version would use one consistent collection method on both sides, sample multiple boards, and run a longer time series.

The full [evidence table](#) is on the site. The [code](#) is on GitHub.

References

- Kalibrr. Job board scraped for both archived and current postings. kalibrr.com
- Internet Archive. Wayback Machine CDX Server API, used to find archived Kalibrr pages. web.archive.org
- Philippine Statistics Authority. Labor Force Survey, “Number of Employed Persons by Major Industry Group,” accessed via the PXWeb statistical API. openstat.psa.gov.ph
- International Monetary Fund (2025). “Artificial Intelligence and the Philippine Labor Market: Mapping Occupational Exposure.” Cited in TAF-AIR’s own project brief. imf.org