

# Workday - How Far Should AI Agents Go?

**Internal ID:** S&F-STR-596

**Topic:** Business Models, Corporate Strategy, Customer Strategy

**Industry:** Application Software

**Source:** Strategy & Finance

**Publication date:** 5 October 2026

**License:** Not specified

**Setting:** August 4, 2026

**Information cutoff:** August 4, 2026, 11:59 p.m. U.S. Eastern Time

**Case convention:** The strategy-team role is hypothetical. The case does not represent a disclosed internal debate, fixed budget choice, or management recommendation.

## Decision assignment

Workday had made its Sana Self-Service Agent available to customers through Flex Credits and announced a partnership that would put Workday agents into Google's Gemini Enterprise environment. These moves offered routes to broader access. Yet public disclosures did not show how often customers used the agents, whether they completed work accurately, what value customers received, or what each task cost Workday to serve. [S2, S3, S4]

You are part of a hypothetical Workday strategy team preparing a recommendation for the next 12-18 months. How should Workday pace AI-agent commercialization across HR and finance workflows? Should it widen availability and task coverage quickly, using Sana and partners to reach more employees? Or should it expand task by task, requiring evidence of accuracy, safe escalation, employee adoption, customer value, and sustainable inference and support costs before adding consequential workflows? A third path could combine broad access to low-risk tasks with tighter gates for actions that affect pay, benefits, financial controls, or employment decisions.

Recommend a rollout and measurement approach. Explain what must be true for it to work, make the strongest case against your recommendation, and identify the evidence that would change your view. The decision is about the pace and scope of the next phase; Workday had already launched Sana. The public record does not identify an internal disagreement over this choice.

## A large installed base, but no AI-specific economics

Workday's fiscal 2027 first quarter ended April 30, 2026. The company reported revenue of \$2.542 billion, including \$2.354 billion of subscription revenue, up 13.5% and 14.3% year over year, respectively. Total subscription revenue backlog was \$27.294 billion, up 10.9%; the portion expected to be recognized in the next twelve months was \$8.806 billion, up 15.5%. Backlog reflects existing-customer expansion, new sales, and renewal timing, among other factors. None of these figures isolates AI-agent demand or attributes growth to Sana. [S1, S2]

Workday reported \$4.353 billion in cash, cash equivalents, and marketable securities and \$616 million of free cash flow for the quarter. Product-development expense was \$705 million, compared with \$663 million a year earlier. These company-wide figures show scale and investment capacity; they do not reveal an AI budget or agent return. Cost of subscription services rose by \$62 million to \$412 million, including a \$47 million increase in third-party

hosted infrastructure. Workday did not break out how much of development or hosting supported Sana. [S1]

Workday guided to a 30.5% non-GAAP operating margin for fiscal 2027. This was a forecast, not a realized margin, and the company did not reconcile the outlook to GAAP because it could not reasonably estimate certain adjustments. The reported Q1 GAAP margin comparison also included a \$166 million restructuring charge in the prior-year quarter. Neither figure establishes the economics of AI agents. [S2]

### Exhibit 1. Selected Workday measures at the cutoff

| Measure                                     | Period                   | Company-reported figure                                | What it does and does not show                                           |
|---------------------------------------------|--------------------------|--------------------------------------------------------|--------------------------------------------------------------------------|
| Subscription revenue                        | Q1 FY2027                | \$2.354 billion; +14.3% year over year                 | Recurring company revenue; not AI-specific                               |
| Total subscription revenue backlog          | April 30, 2026           | \$27.294 billion; +10.9%                               | Contracted revenue not yet recognized; not realized sales or agent usage |
| Twelve-month subscription revenue backlog   | April 30, 2026           | \$8.806 billion; +15.5%                                | Expected near-term recognition; no AI attribution                        |
| Product-development expense                 | Q1 FY2027                | \$705 million; +6.3%                                   | Company-wide expense, not Sana investment                                |
| Third-party hosted-infrastructure increase  | Q1 FY2027 vs. prior year | \$47 million within subscription-service cost increase | Aggregate hosting change; not agent inference cost                       |
| Full-year non-GAAP operating-margin outlook | FY2027                   | 30.5%                                                  | Management forecast, not a reported result or agent margin               |

**Sources and limits:** Workday Q1 FY2027 Form 10-Q and May 21, 2026 earnings release. Growth rates are company-reported. The figures do not isolate AI-related revenue, expense, retention, customer savings, or investment returns. [S1, S2]

### From product access to measured use

In March, Workday introduced Sana from Workday, including Sana Self-Service Agent. Workday said customers could access the Self-Service Agent worldwide through Flex Credits allocated with subscriptions, without a separate license or paywall. Sana Enterprise was available with HCM or Financial Management and through Flex Credits. The launch described more than 300 skills and the ability to find information, take actions, build, and automate workflows. Those are product descriptions and company claims, not independent tests of task accuracy or outcomes. [S3]

Workday emphasized that Sana operates within its permissions, security, and audit framework and can connect work across systems. Such controls may help determine which data an agent can access and what actions it may perform. They do not, by themselves, establish that an answer is correct, that an action is appropriate, or that the workflow saves time or money. Public materials at the cutoff did not report agent-specific completion rates, error rates, human overrides, escalation rates, support burden, inference costs, credit consumption, or

customer-level savings. [S3, S1]

Workday's launch release highlighted one customer, Berner, whose head of AI and analytics said the company made Sana its default AI interface in 40 days, reached 90% adoption, and retired 400 ChatGPT licenses. This is a selected, company-published testimonial. The release did not define the adoption denominator or provide an independent baseline, method, time series, or counterfactual. It may suggest what a successful deployment could look like, but it cannot establish representative customer adoption or causal value. [S3]

Flex Credits make the commercial question less straightforward. Including access may make trials easier and encourage customers to experiment across more tasks. But availability via credits is not the same as actual use, additional revenue, or positive margin. If customers consume included credits, Workday may incur model, infrastructure, and support costs without a separate agent charge. If usage leads to renewal, expansion, or additional paid capacity, the value may appear elsewhere and later. No public source quantified these pathways. [S1, S3]

## Exhibit 2. Announced access and the missing operating evidence

| Publicly described by cutoff                                                                     | Not disclosed in the reviewed sources                                                           |
|--------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| Sana Self-Service Agent available via Flex Credits; no separate license or paywall               | Active customer and employee users; frequency and task mix                                      |
| Workday described 300+ skills and actions across information and workflows                       | Accuracy, completion, override, escalation, and harmful-action rates by task                    |
| One selected customer testimonial reported 90% adoption and 400 retired licenses                 | Definition, denominator, independent validation, representativeness, and counterfactual         |
| Google partnership announced, with Gemini as Sana's default model and model choice for customers | Partner economics, model-level quality and cost, usage, incremental contract value, and margins |
| Workday reported company-wide development and hosting expenses                                   | Sana-specific engineering, inference, customer support, and governance costs                    |

## Distribution can extend reach and create dependence

In May, Workday and Google Cloud announced a strategic partnership. The companies described Workday's Sana Self-Service Agent as available within Gemini Enterprise, so employees could reach Workday workflows through Google's enterprise interface. They also said Gemini would be the default model inside Sana while customers could choose another model. The announcement creates potential distribution and model access; it disclosed no contract economics or independently measured use or benefit. [S4]

The partnership presents a strategic choice about where Workday's control matters. A familiar external interface and a leading model may lower friction, broaden reach, or speed product improvement. At the same time, Workday may depend on an outside model provider and interface for parts of the employee experience. The relationship could help Workday preserve its role in governed HR and finance actions-or give the partner greater influence over discovery,

interaction, and economics. The announcement does not resolve which effect will dominate.

Competition is also moving. Oracle announced 22 Fusion Agentic Applications in March 2026, describing agents integrated into its applications with governance and auditability. These are Oracle's product claims, not independent evidence of deployment, comparative quality, customer value, or profitability. Still, the announcement suggests that customers may evaluate agents as part of broader enterprise suites rather than as standalone assistants. [S5]

## **What research can-and cannot-say**

Two studies suggest that AI effects depend on the task and the way people use the system.

Brynjolfsson, Li, and Raymond studied a GPT-3-based conversational assistant introduced in stages to 5,172 customer-support agents at one software firm. Their preferred estimate was about 0.30 additional successfully resolved chats per hour, roughly 14% of the 2.18-chat mean, with larger productivity gains for less experienced workers. The setting retained humans in control of suggested responses and differs from autonomous HR or finance agents. [A1]

The study's customer outcomes require care. The authors found a significant improvement in sentiment classified by a machine-learning model from support-chat text. Their surveyed Net Promoter Score (NPS) estimate, however, was not statistically significant; the study therefore did not detect a significant average improvement in surveyed customer satisfaction. The authors also observed lower attrition associated with AI access, especially among newer workers, but attrition was measured once per worker, without agent fixed effects. They warn that selection into training could make the association look larger than the causal effect. These outcomes do not establish Workday customer savings, retention, or product economics. [A1]

A preregistered randomized experiment by Dell'Acqua and colleagues assigned 758 consultants to GPT-4 access, GPT-4 plus a prompt-engineering overview, or no AI. In 18 tasks classified as within the tested capability frontier, AI groups completed more tasks, worked faster, and received higher quality scores. In a separate complex strategic task outside that frontier, AI-treated participants were less likely to give the correct answer. The consulting exercise was not a deployed enterprise HR or finance workflow; its estimated effects should not be applied to Workday customers. It does, however, raise a practical question: how will Workday identify tasks where systems perform reliably and route uncertain or consequential work to people? [A2]

The research supports measuring performance by task, skill, and workflow design. It does not supply a Workday productivity forecast, customer ROI, pricing premium, or cost saving. Workday's permissions and audit features may help manage access and actions; academic evidence does not establish that those controls eliminate errors or overreliance.

## **Designing a rollout decision**

A broad rollout could place agents in front of many employees quickly, provide more real-use learning, and help Workday remain relevant as customers adopt external AI interfaces. This path depends on safe task boundaries, effective human escalation, a positive customer experience, and a way to fund model and support costs. Broad availability without observable usage and unit economics could create cost before value.

A task-gated rollout could begin with well-defined, reversible tasks-such as finding policy information or preparing a draft-then add actions only after evidence accumulates. More consequential or ambiguous work could require confirmation or a human decision. This path could protect customer trust and help isolate value by workflow, but it may slow learning, constrain the user experience, or allow competitors and partners to become the default

interface.

These approaches need not be exclusive. Workday could make low-risk assistance broadly available while gating autonomous actions according to consequence and evidence. Students should specify the boundary rather than assume that labels such as “agentic,” “low risk,” or “human in the loop” are self-explanatory.

### Exhibit 3. Candidate measures for a task-level launch gate

| Dimension          | Example measure to define                                                                      | Question for students                                              |
|--------------------|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------|
| Reliability        | Correct completion and material-error rate by task and customer context                        | What counts as correct, and who verifies it?                       |
| Human control      | Escalation, override, and confirmation rates                                                   | Which actions require approval, and when should the agent stop?    |
| Adoption           | Active use, repeat use, task mix, and abandonment                                              | Does “available” lead to sustained use for valued work?            |
| Customer value     | Time-to-completion, error reduction, employee/customer feedback, renewal or expansion evidence | What is the baseline and comparison group?                         |
| Commercial value   | Flex Credit use, overages, paid expansion, renewal outcomes                                    | Does included access create willingness to pay or retention value? |
| Cost to serve      | Model/inference and hosting cost, support burden, governance cost per completed task           | Which costs are incremental, and how do they change with volume?   |
| Partner dependence | Model/provider mix, availability, quality and cost by model, interface-sourced use             | What must Workday retain control over?                             |

The table is a proposed measurement framework, not a list of Workday-reported metrics or targets. Students should propose thresholds, identify who owns each measure, and explain how Workday would respond when a task fails a gate. Any numerical thresholds or financial scenarios are assumptions, not disclosed facts.

## The strategy team’s recommendation

The cutoff is August 4, 2026. Workday’s Q2 FY2027 results for the quarter ending July 31 had not yet been reported. Do not use later company results or product outcomes.

What should Workday do over the next 12-18 months to commercialize AI agents across HR and finance? Set out a rollout sequence, including which tasks could be broadly available, which require task-level evidence or human approval, and how the Google relationship should fit. Name the operating and customer evidence that would justify wider autonomy, and explain what Workday should do if usage grows faster than customer value or if costs remain opaque.

## Sources

- **[S1]** Workday, Inc., Form 10-Q for quarter ended April 30, 2026, filed May 22, 2026. MD&A, “Financial Results Overview,” pp. 26-28; costs, pp. 29-30; AI risks, pp. 48-49. [SEC filing](#).
- **[S2]** Workday, “Workday Announces Fiscal 2027 First Quarter Financial Results,” May 21, 2026. “Fiscal 2027 First Quarter Results,” “Recent Business Highlights,” “Financial Outlook.” [Investor release](#).
- **[S3]** Workday, “Introducing Sana from Workday: Superintelligence for Work That Finds Answers, Takes Action, and Automates Workflows,” March 17, 2026. Product description, customer testimonial and “Availability.” [Investor release](#).
- **[S4]** Workday and Google Cloud, “Expand Strategic Partnership to Bring AI Agents for HR and Finance Into Employees’ Daily Workflows,” May 28, 2026. Partnership scope, “Real Work, Done Where People Already Are,” default Gemini model and model choice. [Joint release](#).
- **[S5]** Oracle, “Oracle Introduces Fusion Agentic Applications,” March 24, 2026. 22 announced applications and governance claims. [Oracle release](#).
- **[A1]** Erik Brynjolfsson, Danielle Li, and Lindsey R. Raymond, “Generative AI at Work,” *Quarterly Journal of Economics* 140, no. 2 (2025): 889-942. DOI: [10.1093/qje/qjae044](#). Sample and design, pp. 890-92, 899-906; productivity Tables I-II, pp. 903-08; NPS Table III, pp. 909-10; sentiment Table IV, p. 931 and Figure X, pp. 932-33; attrition discussion, pp. 934-35.
- **[A2]** Fabrizio Dell’Acqua et al., “Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of Artificial Intelligence on Knowledge Worker Productivity and Quality,” *Organization Science*, Articles in Advance, March 11, 2026. DOI: [10.1287/orsc.2025.21838](#). Study design and sample, pp. 2-4; task results, Tables 4-6, pp. 7-11; outside-frontier task and limitations, pp. 11-18.