Case study · Cloud services · KeyDelta engagement, AI delivery lead · 6 months

MID-MARKET CLOUD SERVICES PROVIDER · 500+ EMPLOYEES · NAME WITHHELD

Internal tickets from 850 to 270 a month in 6 months.

Every department had its own documents, wikis, and tribal knowledge, and employees Slacked a colleague every time they needed an answer. KeyDelta built an AI assistant that gave everyone instant access to HR policies, product specs, sales playbooks, and live business intelligence. The assistant kept earning trust because we fixed the operating layer first: clear ownership of each knowledge source, a governance cadence that keeps answers current, and accountability for accuracy. That discipline separates a living system from an abandoned wiki.

Situation

A mid-market cloud services provider with 500+ employees had a knowledge problem disguised as a people problem. Every department (HR, engineering, product, sales, marketing, finance) had its own documents, wikis, and tribal knowledge. When employees needed answers, they Slacked a colleague, filed an internal ticket, or guessed. The result: constant interruptions, inconsistent answers, and a growing internal support burden that pulled senior people away from high-value work.

- HR fielding 200+ repetitive policy questions per month: PTO, benefits, expense policies, onboarding steps
- Sales reps couldn't find current product specs or competitive positioning without asking product managers
- New hires took weeks to become self-sufficient: tribal knowledge lived in people's heads, not systems
- Marketing and sales using outdated collateral because the latest versions were buried in shared drives
- Finance and ops leaders navigating Power BI dashboards to answer questions that should take 10 seconds

Approach

KeyDelta built a conversational AI assistant that connected to every content source in the company and gave every employee a single place to ask anything.

01 **Content ingestion and knowledge graph**
Indexed content repositories across HR, product, sales, marketing, engineering, and finance. Built a unified knowledge index that understood relationships between documents, policies, products, and processes.

02 **Copilot Studio and Power BI integration**
Deployed on Microsoft Copilot Studio for natural language chat with deep Microsoft 365 integration. Connected Power BI for real-time business intelligence: revenue, service metrics, and operational KPIs, answered from plain-English questions.

03 **AWS Lambda backend intelligence**
Routed complex queries requiring Python-based ML inference and multi-source synthesis to AWS Lambda, using the company's existing AWS infrastructure and ML libraries not available in Azure Functions. Serverless architecture meant costs scaled with usage, not headcount.

04 **Department-specific tuning and rollout**
Tuned response accuracy per department. HR got policy-specific citation and compliance guardrails. Sales got competitive intel with freshness dates. Finance got BI answers with drill-down links. Each department validated before launch.

Vision	Any employee, any question, answered in seconds. No more Slack-a-colleague culture.
Outcomes	Internal ticket volume, answer accuracy, and department time savings measured weekly. The assistant had to earn adoption with quality, not mandates.
Ownership	Each department owned the accuracy of its knowledge source. IT owned the platform. Nobody owned everything: clear lanes, clear accountability.
Cadence	Weekly accuracy reviews, monthly content freshness audits. Stale answers flagged automatically: no more outdated policies in circulation.
Systems	New departments and content sources added without re-architecture. The assistant grew from 3 departments at launch to 6 in 4 months on the same platform, with the same team.

Stack

Microsoft Copilot Studio · Microsoft 365 · Power BI · AWS Lambda

Outcome

Over 6 months, internal tickets fell from 850 a month to 270 and answer time dropped from hours or days to under 8 seconds, with 84% of questions resolved on the first ask. Employee NPS rose from 32 to 61, customer CSAT from 4.0 to 4.4, and the company avoided $420K a year in headcount. Builds like this now ship under the KeyDelta Method and stay current under AAMP.

−68%

Internal tickets

From 850 to 270 a month over 6 months

<8 sec

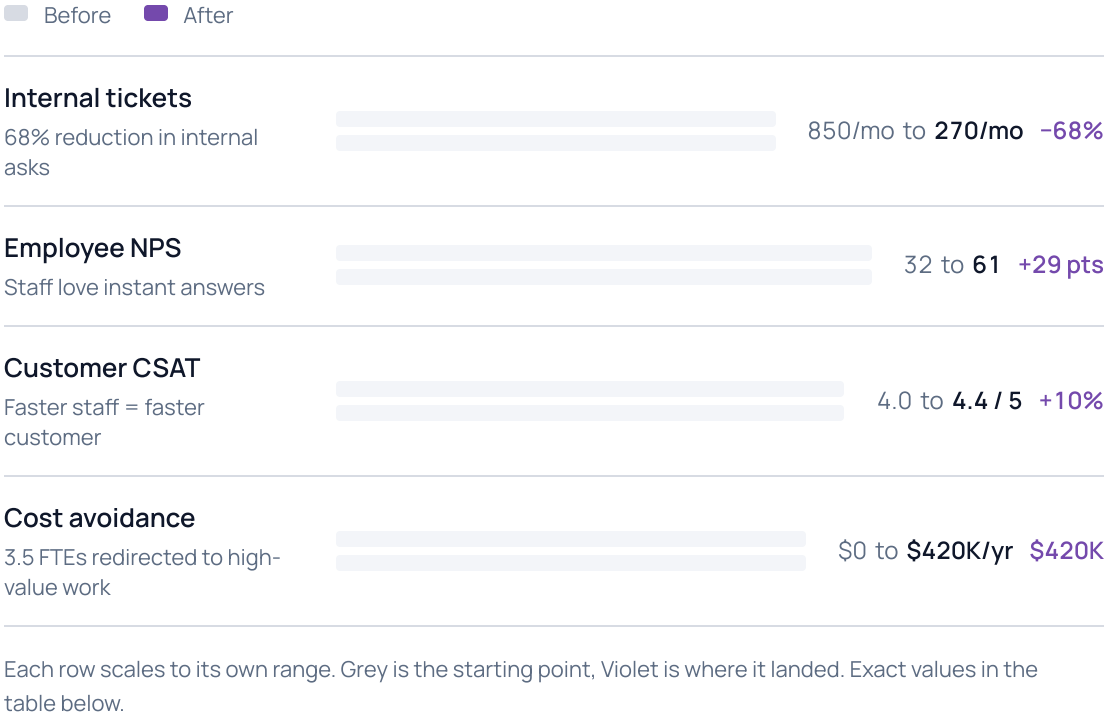
Answer time

From hours or days, measured at month 6

4.4x

AI ROI

In 6 months, KeyDelta engagement



Each row scales to its own range. Grey is the starting point, Violet is where it landed. Exact values in the table below.

MEASURE	BEFORE	AFTER	CHANGE
Internal tickets. 68% reduction in internal asks	850/mo	270/mo	−68%
Answer time. From waiting to answers	Hours/Days	‹ 8 sec	Instant
First-ask accuracy. Resolved without follow-up	N/A	84%	84%
Employee NPS. Staff love instant answers	32	61	+29 pts
Customer CSAT. Faster staff = faster customer	4.0	4.4 / 5	+10%
Cost avoidance. 3.5 FTEs redirected to high-value work	$0	$420K/yr	$420K

"We had 500 people constantly interrupting each other for answers that already existed somewhere in our systems. The AI assistant didn't replace anyone, it freed everyone. Employee NPS jumped 29 points. New hires onboarded in 8 days instead of three weeks. We avoided $420K in headcount by redirecting people from answering repetitive questions to actual high-value work. Our customers noticed too, response times dropped because our people had answers at their fingertips."

KeyDelta operating lead

$125M+ revenue impact, delivered to clients across 30+ transformations. Measurable results by day 30, typical across embedded engagements. 3.8 to 5.1x AI ROI in 6 to 9 months, delivered to clients. Client anonymized; details shared with written approval.

Book a 30-Minute Call

Builds like this ship through our **AI agent development** and stay current under our **managed AI services**.