Customer support assistant study

IMPLEMENTED 2020–2021 · TECHNOLOGY AND SOFTWARE · US · BAND D · PROVED
VERSION 1.0 · PUBLISHED 15 NOVEMBER 2026 · UPDATED 15 NOVEMBER 2026 · FACTS CHECKED 15 NOVEMBER 2026

5,172 customer support agents at a software firm were given an AI assistant in stages; the largest gains went to the least experienced.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

AI used: Generative AI: a GPT model fine-tuned on past support chats, suggesting replies to agents.

Implemented: 2020–2021. Grade: band D, Task support, Proved.

Headline result, as reported: 15% more issues resolved per hour on average, in a peer-reviewed study

1 · The case

The facts, as reported by the 2 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

A Fortune 500 software firm (Brynjolfsson, Li and Raymond). A Fortune 500 enterprise software company selling business process software to small and medium-sized US businesses. Its chat-based technical support agents, employed directly and through outsourcing firms, were mostly outside the US, mainly in the Philippines. Canvas block: 5. People and skills. Also 4. Operations and process, 6. Data and technology.

Business challenge

The challenge. Not reported in the firm's own terms. The authors describe the setting: newer support agents need significant training and take time to become productive, and industry estimates put contact centre staff turnover at 60% a year.

How it was addressed. The firm bought an AI assistant from a customer service software vendor. Built on a GPT model fine-tuned on past support chats and weighted towards top performers, it suggested replies and links to internal documentation that only the agent could see. After a

Who implemented it

Customer service software firm (not named) (Software vendor): Supplied the assistant.
OpenAI (Model or cloud provider): Supplied the underlying GPT model.

Technology (SEAM pillar E)

SEAM types: T3 API integrations, T5 Custom / fine-tuned models. Deployment: Vendor SaaS.
Vendors and products: OpenAI, GPT (version not named).

Where it was built in: Bought from an AI customer service software firm and shown to support agents in the chat window, suggesting replies and links to internal technical documentation in real time.

What changed

Before: agents answered customer chats unaided. After: agents saw real-time suggested replies and documentation links, and chose whether to use, change or ignore them. [1] [2]

What it cost

Not reported. The authors say the firm set a limited budget and a contractual cap on licences.

What it saved

15% more issues resolved per hour on average, by the authors' analysis of 3 million chats. Less experienced and lower-skilled agents gained most in speed and quality; the most experienced saw small gains in speed and small falls in quality. Requests to speak to a manager fell by almost 25%, and attrition among agents with under six months' experience fell by about 40%. No economically significant change in customer satisfaction. Achieved cash saving: Not reported. [1] [2]

Governance and control

Suggestions were shown only to agents, who had full discretion over whether to use them; the authors say this reduces the chance that incorrect outputs reach customers. Agents had no further training after onboarding. Accountable owner and policy: Not reported.

S2 · Accountable owner · Not reported
Not reported.

A1 · Safe use · Not reported
Not reported.

A2 · Human in the loop · Evidenced
Output was shown only to the agent, who had full discretion over whether to use it, the study reports. [1]

A3 · Monitoring and assurance · Evidenced
The AI firm trained the model to favour empathy and limit unprofessional language, and the researchers measured resolutions per hour, customer sentiment and requests to speak to a manager. [1]

A4 · Vendor AI · Not reported
Not reported.

A5 · Agent controls · Not reported
No agents described.

A6 · Standards and regulation · Not reported
Not reported.

What worked, what did not, what is unproven

WORKED

– Productivity rose 15% on average, measured as issues resolved per hour.

– Less experienced and lower-skilled agents improved in both speed and quality.

– Customers were more polite and less likely to ask for a manager.

– Attrition fell, most among newer agents.

DID NOT

– The most experienced and highest-skilled agents saw small falls in quality.

– Customer satisfaction did not change in an economically significant way.

UNPROVEN

– Effects on wages, overall demand for labour and hiring, which the data do not cover, by the authors' account.

– Whether the results hold beyond one firm and the medium run.

– Whether the model keeps improving if top agents contribute fewer original answers, a risk the authors raise.

Timeline

Aug 2020 Pilot with about 50 agents, half randomly given the tool [1]

Oct 2020 Rollout under way, with about 5% of agents having access [1]

Jan 2021 About 70% of agents have access [1]

Apr 2023 Findings released as a working paper [2]

2025 Published in the Quarterly Journal of Economics [2]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: Not applicable. Grade: D, Task support. Proof: Proved. Level of change: Task support to Process change.

Task support moving towards process change (Level 1 to 2, band D). Agents kept the same job and workflow; the assistant shaped how each chat was handled.

An assistant to agents, not a redesigned process.

The SEAM view and lessons learnt

The gain came from how the tool met people: it spread the habits of top performers to newer staff, and agents kept the final say. That pattern transfers well to a large service team. The authors' warning matters too: if top staff lean on the tool, the expertise it learns from may thin out over time.

2. Keep people in control of what reaches customers.
3. Protect the expert input the system learns from.

Evidence

Score 2 of 4. Independent: academic researchers using the firm's data; peer reviewed in the Quarterly Journal of Economics.

Evaluation method. Comparison group: a staggered rollout analysed against agents not yet given access, after a small randomised pilot; peer reviewed.

Sources disagree. The arXiv version (November 2024) reports 5,172 agents and a 15% average gain; the NBER page summary reports 5,179 agents, a 14% gain and a 34% gain for novice and low-skilled workers.

Status. Completed: a study of a 2020–2021 rollout, published in 2025.

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/customer-support-assistant-study

[1] Brynjolfsson, Li and Raymond, "Generative AI at Work"
https://arxiv.org/pdf/2304.11771v2

[2] NBER: working paper and published version in the Quarterly Journal of Economics
https://www.nber.org/papers/w31161

Version 1.0 · Published 15 November 2026 · Updated 15 November 2026 · Facts checked against sources 15 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/customer-support-assistant-study

VERSION HISTORY

v1.0 15 November 2026 First published