True Fulfillment Coaching

Unlock Potential. Achieve Clarity. Create Bold Change.

PM PORTFOLIO · WORKED CASE STUDY

Smart Draft

ILLUSTRATIVE EXAMPLE · FICTIONAL PRODUCT & DATA · 2026

AI PRODUCT · B2B SAAS · HUMAN-IN-THE-LOOP

Smart Draft: An AI Reply Assistant for Support Agents

How a support team moved from copy-paste exhaustion to a 32% faster first reply — without letting AI speak to a customer unsupervised.

The Challenge

Support agents were spending roughly forty percent of every ticket re-typing near-identical answers. One agent put it plainly: *"I answer the same billing question thirty times a day — I'm a copy-paste machine, not a problem-solver."*

The business felt it too. Median first-reply time stood at 2h18m — the single worst-scoring driver of customer satisfaction — and rising ticket volume had produced a request to hire six more agents, roughly $540K a year. Reducing handle time had become a board-level cost lever.

Research & Insights

Nine agent interviews, six shadowed live shifts, four thousand tickets clustered by intent, and a review of two competitor assistants produced three findings that shaped everything downstream:

- **Sixty percent of tickets fall into twelve repetitive intents** — a narrow, high-volume target where AI can actually be accurate.
- **Agents fear wrong AI answers reaching customers.** Seven of nine raised it unprompted. Trust had to be designed in, not bolted on.
- **Agents already personalize every reply.** Auto-send was the wrong model; draft-to-edit was the right one.

The Problem, Stated Once

Support agents need a way to answer common tickets faster without sacrificing accuracy or their personal touch — but they do not trust AI to talk to customers unsupervised. The job to be done: *when a routine ticket comes in, help me produce a correct, on-brand reply in seconds that I can quickly personalize and send.*

ROLE

Sole PM
Concept through launch

TEAM

3 engineers
1 designer
1 data scientist

TIMEFRAME

10 weeks
Before renewal season

CONTEXT

Mature B2B support platform · net-new AI feature · strict data-privacy constraints

OUTCOME SNAPSHOT

First-reply time **−32%** · CSAT **+6 pts** · unsafe-suggestion rate held under **2%**

PRIORITIZATION · PRODUCT JUDGMENT

Deciding What Not to Build

The riskiest assumptions were tested first; the most tempting option was the one deliberately killed.

Assumptions, Tested Before Building

ASSUMPTION	RISK	HOW IT WAS TESTED
The model can exceed 90% factual accuracy on the top twelve intents	High	Offline evaluation on 500 labeled tickets — 92% before launch
Agents will trust a draft-to-edit model	High	Prototype test with eight agents — strong preference over auto-send
Time saved will not simply be spent editing	Medium	Timed task study — a net 90 seconds saved per assisted ticket

The Decision

- **Build now — draft-to-edit assistant on the top twelve intents.** Best risk-to-reward; keeps humans in control.
- **Cut — the fully automated reply bot,** despite executive excitement. One unsupervised wrong answer was an unacceptable trust and brand risk, and the research showed agents would not adopt it.
- **Later — knowledge-base search sidebar.** Useful, but smaller savings.
- **Cut — real-time voice assist.** Out of scope for a ten-week window.

Constraining scope to draft-to-edit is precisely what made the launch both safe and fast.

> *Nothing reaches a customer without a human pressing send — that single constraint was the product strategy.*

The Solution

Smart Draft watches an open ticket, classifies its intent, and generates a suggested reply in a side panel with its knowledge-base sources cited. The agent edits inline and sends. When the model's confidence falls below threshold, it shows nothing rather than a bad draft — the feature fails safe, not silent.

The flow: ticket opens → intent detected → draft and sources appear → agent edits → agent sends → edit distance is logged to improve the model.

Feedback & What Changed

- Beta agents distrusted drafts with no rationale — **inline source citations** were added to every draft.
- A confident-but-wrong draft proved worse than none — a **confidence threshold** now suppresses weak drafts entirely.
- Agents wanted to teach the system — a one-click **"bad draft" signal** now feeds the evaluation set.

RESULTS · SAFETY · REFLECTION

The Outcome

Every target defined before the build; every number reported after it — including the ones that could have missed.

Measured Against Its Own Targets

METRIC	TARGET	ACTUAL
Median first-reply time	−25%	**−32%**
Draft acceptance rate (sent with minor edits or fewer)	≥ 50%	**58%**
CSAT, rolling 30-day	+3 pts	**+6 pts**
Unsafe or incorrect suggestion rate	< 3%	**1.7%**
Escalation / reopen rate	No increase	**−0.4 pts**

A three-week A/B test across eighty agents confirmed the result; the feature then rolled out to all 210. Projected impact: four of six planned hires deferred, roughly $360K a year. One agent's verdict: *"It handles the boring sixty percent so I can actually help on the hard tickets."*

Quality, Safety & Trust

- **Evaluation.** A 92% factual-accuracy bar on 500 labeled tickets to ship; weekly online evaluation of sampled live drafts thereafter.
- **Human in the loop.** A hard rule — no AI text reaches a customer without an agent editing and pressing send.
- **Guardrails.** PII, refund, and legal intents route to humans with no AI draft at all; low-confidence cases fall back to the normal manual flow.
- **Trust by design.** Every draft cites its sources; every draft can be flagged with one click; every edit is logged.

Tradeoffs & Reflection

Coverage was traded for safety: twelve intents at launch left roughly forty percent of tickets unassisted — a price worth paying to protect trust. The remaining risks are model drift, mitigated by the weekly evaluation job, and over-reliance dulling agent judgment. With hindsight: instrument edit-distance from day one, and bring the QA team in earlier to co-own the evaluation set.