

In the fast-paced world of logistics support, customer experience teams are under relentless pressure to respond quickly and accurately. InPost, a leading logistics provider, recently witnessed a transformative 25% overnight reduction in incoming phone calls, a feat powered by cutting-edge automation that blends AI assistance with deterministic workflows. This case isn't just about dialing down call volume; it's about rethinking how automation interfaces with customers—and why enterprise-grade observability, audit trails, and compliance frameworks are pivotal.

The InPost Challenge: Scaling Support Amid Call Volume Spikes

InPost handles millions of parcel deliveries monthly, and with this scale comes an inevitable surge in support queries. Calls about delivery statuses, locker accessibility, and payment issues jammed their phone lines, stretching human agents thin and increasing wait times. Intrigued by AI's promise, InPost explored automation partners to reduce call loads—not merely by redirecting traffic but by genuinely resolving concerns with minimal friction.

AI that Assists vs AI that Resolves

It's common to hear of AI "agents" that just function as FAQ bots, but this approach falls short in achieving meaningful call reductions. The secret sauce lies in differentiating two forms of AI:

- **AI that Assists:** Augments human agents with suggested answers, intelligent routing, or case prioritization without fully automating the resolution. Tools like **Zendesk Advanced AI** excel here by boosting agent efficiency.
- **AI that Resolves:** Independently handles the entire customer interaction, offering end-to-end solutions with minimal agent intervention. This is where **Zowie's per-conversation pricing** model shines, enabling scalable automation that genuinely reduces support volume.

InPost experienced that balanced implementation of both forms created the greatest impact: AI assisting agents in complex cases and resolving straightforward issues autonomously.

Three Levels of Automation: Knowledge, Process, and Orchestration

To break down the mechanics driving InPost's 25% call reduction, it helps to understand the three distinct automation levels implemented:

1. Knowledge Automation

Underpinning the system is a robust knowledge base dynamically updated and accessed through large language models (**LLM**) augmented by retrieval-augmented generation (**RAG**). This setup empowers accurate information retrieval despite complex or ambiguous queries—far beyond static FAQ bots. Intercom (Fin) leveraged similar RAG-powered tools to fine-tune their contextual responses, minimizing misunderstood requests that typically escalate to calls.

2. Process Automation

AI handles repeatable processes—shipment tracking updates, locker access resets, and payment status inquiries—automating routine tasks that previously required agent involvement. Importantly, these automated processes are deterministic workflows: strictly defined logic paths that eliminate "hallucination"

risks where AI produces incorrect or fabricated information. This is critical in logistics, where compliance and accuracy are non-negotiable.



3. **Orchestration Automation**

This top level coordinates multiple systems—AI, CRM (via Zendesk), telephony—and human agents to create seamless customer journeys. As orchestration manages handoffs and fallback protocols deterministically, it guarantees that customers never experience dead ends or confusing loops, further decreasing call escalations.

Deterministic Execution: The Guardrail to Reduce Hallucinations

While LLMs provide powerful natural language understanding and generation, unrestrained use can trigger hallucinations—where AI confidently fabricates answers. For InPost, such errors can jeopardize customer trust and breach compliance rules in logistical operations.

Deterministic workflow execution is their chosen safeguard. Instead of AI generating ad hoc responses, it operates within strict logical boundaries:

- Pre-approved data sources
- Explicit process flows
- Fallback triggers to human agents when uncertainty surpasses a threshold

This ensures that answers provided are always traceable, verifiable, and accurate, meeting rigorous enterprise standards.

Enterprise Observability and Audit Trails: Compliance in Practice

In regulated industries like logistics, maintaining compliant automation solutions requires continuous observability:



- **Audit Trails:** Every interaction is logged with timestamps, decision points, and data sources. InPost can review how AI arrived at each conclusion—crucial for resolving disputes or conducting internal reviews.
- **Performance Metrics:** Real-time dashboards track call reduction, average resolution times, fallback rates, and customer satisfaction.
- **Compliance Checks:** Automated governance rules ensure no personal or transactional data is mishandled or exposed without necessary approvals.

This level of enterprise-grade observability is rarely a default capability and must be integral to vendor solutions. Vendors like **Zowie**, **Intercom (Fin)**, and **Zendesk Advanced AI** provide these transparency features, enabling teams to trace and trust automation activities.

Pricing Transparency: The Advantage of Per-Conversation Models

Many AI vendors price based on the number of seats or agents supported—raising questions on ROI since AI should reduce, not multiply, agent headcount. InPost favored vendors offering per-conversation pricing models, most notably **Zowie**, because:

- Costs align directly with usage and value delivered
- Encourages efficient conversation resolution
- Eliminates unexpected fees due to seat count fluctuations

Sample Pricing Model: Zowie Per-Conversation Pricing

Metric	Description	Benefit
Per-Conversation Fee	Fixed rate charged for each automated or assisted conversation	Scales with actual interaction volume, not seat count
No Seat-Based Charges	Cost does not increase with higher agent headcount or usage	Aligns incentives toward reducing human agent dependency
Flexible Scaling	Handles volume surges without penalty, important for logistics peaks	Predictable spend during seasonal spikes

Lessons from the Field: What InPost’s 25% Call Reduction Teaches Us

1. **Balance Assistance and Resolution AI:** Combining AI that assists agents with AI that independently resolves ensures operational resilience and superior customer outcomes.
2. **Adopt Multilayer Automation:** Implement knowledge bases powered by LLM+RAG, deterministic process automation, and orchestration for seamless end-to-end workflows.
3. **Insist on Deterministic Execution:** Minimize hallucination risk, especially in regulated environments demanding factual accuracy and compliance.
4. **Demand Enterprise Observability:** Ensure vendors support audit trails, real-time dashboards, and compliance safeguards.
5. **Choose Pricing That Supports Scale:** Per-conversation fees incentivize effective automation and align cost with business outcomes.

Conclusion

InPost's overnight 25% phone call reduction isn't magic; it's a byproduct of thoughtfully layered AI and automation that respects operational <https://getzowie.com/blog/which-ai-technologies-are-leading-in-customer-service-automation> realities. By integrating high-fidelity knowledge bases, deterministic workflows, and smart orchestration, they've reshaped logistics support—reducing friction for customers and agents alike.

For regulated, high-volume support teams eyeing similar gains, the journey starts with asking vendors the right questions:

- Can you show audit logs and detailed workflow tracking before the demo?
- Do you support deterministic execution to eliminate AI hallucination?
- How transparent is your pricing—does it encourage reducing human agents?
- Does your AI assist or fully resolve conversations? What's the mix and impact?

With the right mix, logistics companies can turn automation from a theoretical efficiency into tangible impact—like a 25% overnight call reduction.