



Support Assistant

OUR APPROACH TO CUSTOMER 1ST, STARTS WITH CUSTOMER ZERO

Increased customer and support efficiency with Support Assistant

Elastic's Support Assistant empowers support engineers to efficiently respond to customer support issues and enables customers to self-serve via a generative AI chatbot experience. It delivers business impact by freeing up engineers' time to focus on more complex projects and helping customers quickly answer their queries.

586

case deflections in first month

4 months

payback period

400

returning customer users



Customer success as customer zero

Modern enterprises realize the promise of generative AI for reshaping customer experiences and driving greater efficiencies. Over 90%¹ of C-suite executives believe automating manual workflows for customer support, using AI assistants for day-to-day tasks, and providing self-service support experiences can deliver tangible business impacts. To better serve our current customers and provide best practices to future customers, Elastic led by example and built a generative AI assistant using [Elasticsearch](#) to improve engineer workflows and enhance customer experiences.

Our field engineering team's objective? Build a scalable, secure, and accurate chat experience for engineers and customers that uses our various internal and external data sources to deliver highly relevant responses. Today, Elastic's 200+ internal support engineers use the Support Assistant to resolve customer cases. Within the first month of launching the Support Assistant for customers, we've served 1,309 customers and deflected 586 cases.

“Using generative AI for support has been top of mind since the launch of ChatGPT. In the past few months, this idea has become a reality. As a support team, we deal with thousands of cases every week, and we are process-heavy and data-rich. Within the first month of the launch of Support Assistant, my team has already seen capacity gains because it is easy to refine our responses and find information to support customer queries.”

Julie Rudd, VP Global Customer Support, Elastic

1. “Solving business challenges with data & AI” Elastic, 2024

The challenge

Handling customer queries at Elastic faced many obstacles like:

- **Reactive service:** Previously, support was a reactive process. A customer opened a case, and Elastic reacted to resolve the query.
- **Wait time:** Either the customer or the support engineer had to wait for answers, leading to a time-consuming process.
- **Customer expectations:** There is high pressure to deliver fast and efficient service to customers along with a growing customer demand for self-service to speed up time-to-resolution.
- **Unique deployments:** Often serving multiple customers simultaneously using different product versions and deployment models, support engineers needed to meet each customer's distinct needs.
- **Information overload:** Support engineers and customers spent too much time looking for relevant answers from sources across the knowledge base (including 300K+ product documentation, internal knowledge base articles, blogs, previous customer cases, communities, etc.).
- **Case fatigue:** Support engineers spent significant time processing case summaries for escalations or transitioning cases from one engineer to another.

“Support Services are typically reactionary and manual and include some wait time between interactions. For example, if you’re trying to figure out what’s wrong with your phone or something broken in your home, the first place you go is Google or now ChatGPT. The technical industry is no different. All customers, including ours, want instant self-service. They want to search, find, fix, and move on.”

Julie Rudd, VP Global Customer Support, Elastic

The solution

From support engineers to support heroes with Elastic's Support Assistant

Our traditional users are developers, site reliability engineers, security analysts, and business analysts seeking answers and insights for business questions. Most customers try to find relevant answers first before reaching out to support. Elastic's Support Assistant delivers a generative AI chatbot experience, built on Elastic Search, to meet self-service expectations by quickly answering natural language questions.

In the Support Assistant, users can pose a natural language question and receive an easy-to-understand answer and recommendation in seconds. Internally, Elastic's support engineers use the Support Assistant to query internal documentation to enrich their case responses. Externally, customers use the exact same tool for a self-service RAG Generative AI experience. Rooted in knowledge-centered support (KCS), the Support Assistant also allows support engineers to deliver real-time knowledge and answers to customers looking for them by easily drafting new knowledge documentation. This looping mechanism allows Support Assistant to integrate knowledge creation and maintenance in real time, continuously increasing quality and relevance of the knowledge base.

Elastic's Support Assistant maturity and evolution

The Support Assistant — designed to answer technical questions — requires an underlying architecture that delivers accurate and relevant answers as technically deep and broad as Elastic's solutions. But to build the current generative AI experience, the team went on a search maturity journey.

The journey to build the Support Assistant started three years ago with simple keyword-based search experience on the support portal. Customers and support engineers could search the support portal for links to relevant documentation and resources based on their queries. However, textual search is unable to derive semantic meaning. Thus we added more complex vector search capabilities enabling Elasticsearch to derive a semantic interpretation of the user's query and deliver more relevant results. This is especially important in customer support when customers may not always be aware of relevant keywords relating to their issue.

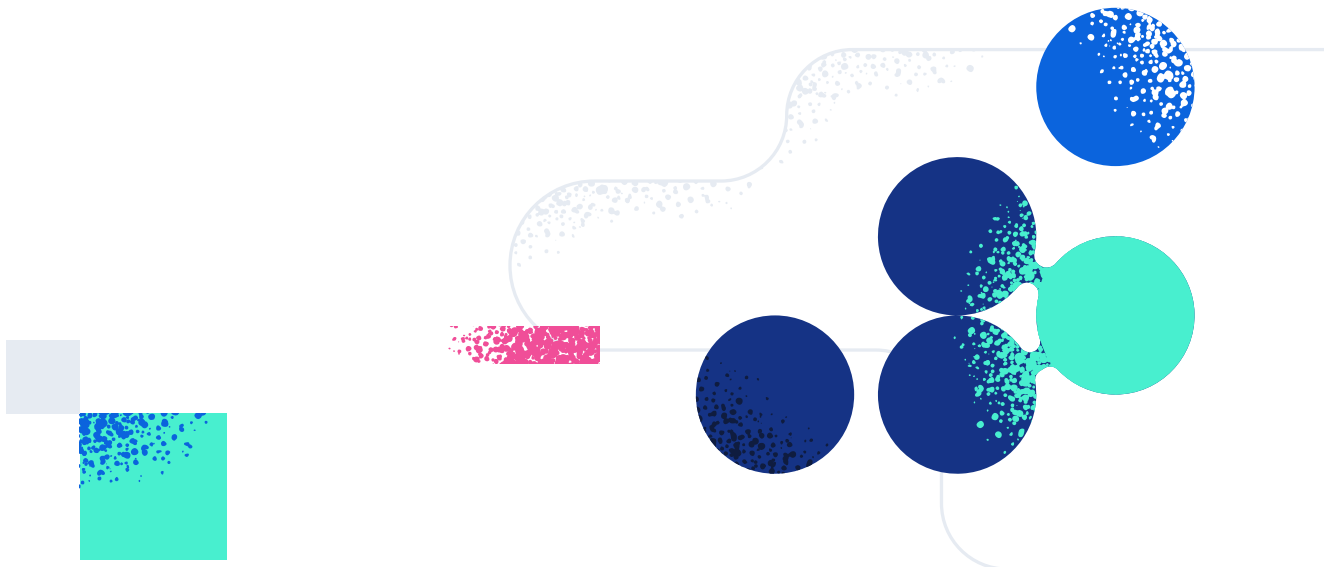
For the most recent evolution, we built a Search AI chat experience using a retrieval augmented generation (RAG) architecture and a large language model (LLM). We launched it internally for testing accuracy and landed on a RAG-based approach for several reasons:

1. It's easy to use with unstructured data and immediately incorporates real-time information into the knowledge base via the KCS methodology.
2. It restricts access to information with role-based access control (RBAC) and document-level security.
3. It requires less maintenance and effort than fine-tuning our own LLM.

At each phase, relevancy tuning and testing ensured accuracy before released it to customers. To monitor the performance and availability of the Support Assistant, we used [Elastic Observability](#) for logging and application performance monitoring (APM).

“When we started with keyword search, generative AI wasn’t even an option. The advantage now is that using Elastic, you can start with semantic search, combine it with a generative AI model, and build a working solution in a few weeks. Elastic’s search capabilities combined with AI advancements, allowed the team to see results and iterate faster on our self-service initiative”.

Chris Blaisure, Sr. Director, Field Technology, Elastic



Use cases

Providing users with real-time answers

The initial use case for Support Assistant was to deliver self-service experiences to engineers and customers. Moving a majority of customers to a self-service experience has allowed our Support Engineers to focus on high-priority, complex cases, and strategic projects. However, there are several other use cases and benefits to this generative AI chatbot:

Internal support experience	External customer experience
Support engineers use Support Assistant to find relevant information, sources, and answers for ongoing support cases.	Customers use Support Assistant via the Elastic Support Hub for quick, self-service answers to product queries using natural language.
<ul style="list-style-type: none">• Drafting initial reply: Streamlines process of drafting initial responses with quick access to relevant answers in professional written language.• Augmenting case summaries: Quickly builds case summaries for escalations or case transitions, enhancing efficiency, effectiveness, and issue resolution.• Knowledge documentation drafter: Efficiently creates and updates documentation, expanding support knowledge base for KCS and facilitating quicker problem resolution for engineers and customers via real-time updates.• Onboarding and enablement: Fast and easy access to product and feature information for support and employee onboarding, improving product enablement and knowledge.	<ul style="list-style-type: none">• Troubleshooting configurations: Tailors guidance for troubleshooting based on deployment or Elastic version configurations, saving customer time.• Performance tuning and upgrades: Easily optimizes performance and upgrades with step-by-step guidance based on specific deployment needs.• Security and compliance: Provides immediate suggestions for securing an Elastic deployment, including setting up document and RBAC as well as resources on compliance requirements (GDPR, HIPAA, etc.)• Custom use cases and integrations: Build custom applications with information for code snippets, integrations, and relevant examples for specific needs.

“The original intent was a self-service generative AI customer experience that would deflect the volume of easy support cases and allow my team to focus on more complex cases that required human intervention. But then we benefited from many secondary use cases, allowing our support team to scale as our customer base grows”.

Julie Rudd, VP Global Customer Support, Elastic

The results

Initial metrics from the first month of launch

The business impact of the Support Assistant has just started, and Elastic is already capitalizing on efficiencies and recouping its investment.

586

case deflections in first month

4 months

payback period

400

returning customer users

Since launching the Support Assistant, 4% of our daily customers navigating the Support Hub have already started using it for support queries, resulting in nearly 600 case deflections. Taking into account the cost of running the support assistant, including hosting the LLM, and the labor costs associated with building this solution, the support team expects a return on investment of this given solution within four months².

Additional benefits

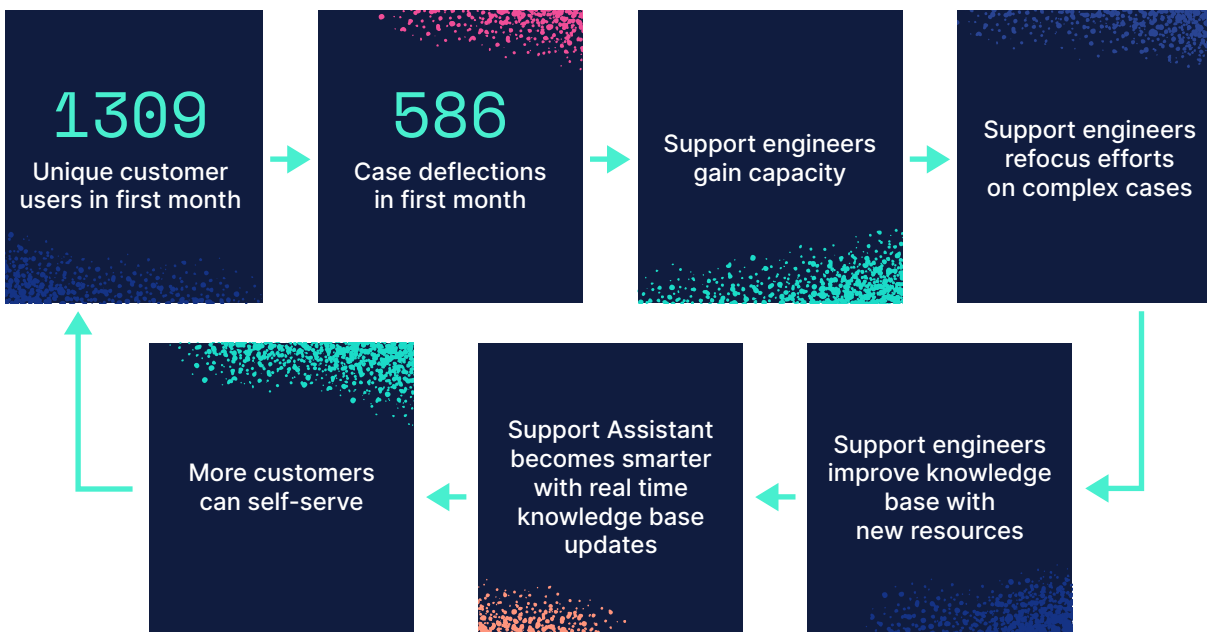
There are several other benefits of the Support Assistant:

- **Mean time to close:** With the deflection of easy support cases, the team is increasingly spending more time on complex cases. In the first month since launch, we've seen an **11% reduction in first contact close rate**. Though this leads to increased mean time to close, the support organization is investing its human capital on the most complex cases that matter the most.
- **Customer adoption and satisfaction:** With 400 repeat users in the first month of launch, the team expects adoption to increase as more customers learn and interact with the Support Assistant. As more customers use the tool, Support Engineers will have more time to spend on complex cases, leading to increased quality of support engagements and improved customer satisfaction from the overall support experience.
- **Improving knowledge-centered service (KCS):** As more complex cases arise, support engineers will continue to build and add relevant resources and content to the knowledge base in real time. This will make it easier for customers to self-service and find relevant answers, eventually speeding up time to resolution for both self-service and incoming cases.
- **Accelerating onboarding:** New hires are using the Support Assistant for learning and development, taking pressure off of senior engineers and accelerating the time it takes for new employees to become fully productive in their roles.

2. This forecasted calculation only considers resource capacity gains through productivity recapture of the 586 case deflections per month for "easy cases." Although increased customer adoption will lead to more case deflections over time, this payback period calculation still does not include any growth in case deflections over the first year.

The gift that keeps on giving

As the Support Assistant adoption grows and support engineers see more capacity gains, the tool will continue to deliver more relevant responses via self-service through an improved knowledge base. This constant feedback loop over time will allow the support function to grow and scale in new ways.



“Since ChatGPT showed up in the support industry, there’s a preference for instant answers. Elastic customers expect that from us because of who we are and what we do. Budgets are constantly under pressure. How do I scale support with a growing business, make sure I’m getting the maximum value from my support budget, and reduce my cost envelope? I will hire data scientists and technology experts to build technology that delivers a best-in-class service to our customers.”

Julie Rudd, VP Global Customer Support, Elastic

What's next

We're consistently building new use cases from the Support Assistant experience.

- 1. Enriching support cases:** To save more time and create consistency, the team is currently working on enriching support cases with additional metadata (product, component, deployment type, root cause, solution codes, etc.) from the case content and other support source systems to make it easier for support engineers to work on incoming cases.
- 2. Support Assistant case-aware:** With more context, customers can ask questions about their past and current cases, delivering more relevant responses.
- 3. AutoOps integration:** Integrating Support Assistant with [AutoOps](#), Elastic Cloud customers will be able to ask questions related to the health of their deployment, enabling faster self-service resolution.
- 4. Automating knowledge creation:** Support engineers can automatically generate knowledge articles in a specific output template from relevant case context, use case notes and Slack threads to make them instantly available for search.
- 5. Multi-language support:** Today, support engineers and customers can ask the Support Assistant to translate existing knowledge and answers into multiple languages. We are working to provide a more personalized, multilingual experience for Elastic customers via the customer-facing Support Assistant.

"I'm already working with Chris to augment further processes within the Support Assistant to improve the productivity of the support team. We want to let genAI populate relevant information from source systems to save time and create better data consistency around the organization. This will not only allow my team to become more efficient but also create business insights for my day-to-day decision-making."

Julie Rudd, VP Global Customer Support, Elastic

As customer zero, Elastic has the expertise, experience, and technology to build generative AI experiences.

[Learn more about how you can do the same with generative AI](#) →