



The Alibaba Challenge: How to Effectively Engage with a Billion Customers



©shutterstock.com/Oleksandra Bezverkha

Yitong Wang
Alibaba

Ofer Mintz
University of Technology
Sydney and Tel Aviv University

Depin Chen
Alibaba

Kehan Chen
Alibaba

Alibaba conducts millions of transactions involving billions of dollars each year. Yitong Wang, Ofer Mintz, Depin Chen, and Kehan Chen describe how the company uses AI-powered chatbots to complement its human resources so as to maximize engagement with its millions of customers and with its vendors.

Alibaba Group, China's largest e-commerce company and one of the world's biggest companies, has nearly one billion annual active Chinese consumers who daily participate in hundreds of millions of transactions through its Taobao e-commerce platform.

During Alibaba's busiest shopping period in 2021 - Taobao's Double 11 shopping festival - its platforms facilitated 540 billion RMB (US \$85 billion) in transactions over a two-day period. The sheer number of customer interactions in Taobao transactions makes it difficult for Alibaba to keep up with proactively engaging with its customers.

Further, Alibaba needs to resolve several million customer service queries each day. The queries can come from end-user consumers or business merchants in Alibaba's two-sided platform. Consumers may have questions about the Alibaba platforms or be unsatisfied with products they purchased from merchants on the platform. Merchants could have questions about end-user consumers or about the platforms.

Taken together, Alibaba faces resource restrictions that limit its

ability to engage with customers solely by human service interactions. Since 2015, Alibaba's response has been to implement artificial intelligence (AI) chatbots, supplemented by human service interactions, to both proactively and reactively engage with its customers.

Alibaba uses AI chatbots to handle customer engagement for more than two million daily sessions and over 10 million lines of daily conversation.

Today, Alibaba uses AI chatbots to handle customer engagement for more than two million daily sessions and over ten million lines of daily conversation on Taobao's two-sided platform, representing about 75 percent of Alibaba's online and 40 percent of phone hotline consultations.

Not only has the use of AI chatbots raised customer satisfaction by 25 percent, based on initial results, it has saved the company

more than one billion RMB annually (~US \$150 million) by employing AI instead of human contact center agents.

Alibaba's AI chatbots

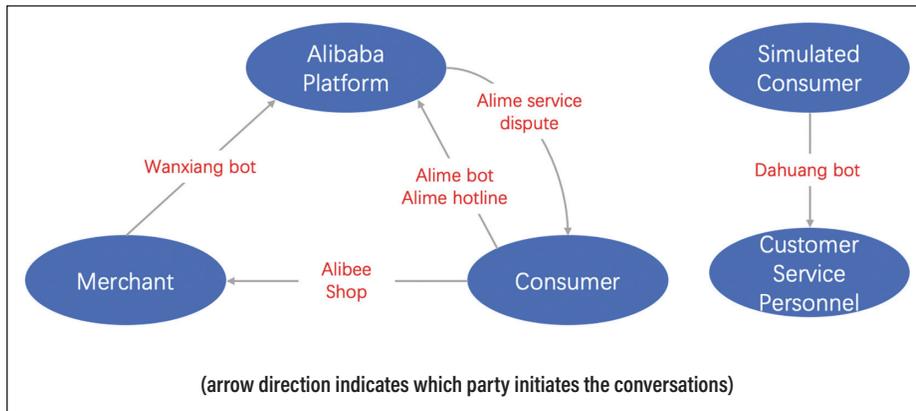
Alibaba employs five AI chatbots to cater to the heterogenous demands from customers on Taobao's two-sided platform.

- The Wanxiang-bot assists merchants on the Taobao platform, performing tasks such as resolving questions about the platform's rules, activities, and service issues.
- The Alibee Shop bot assists merchants with end-user consumer interactions on Taobao, helping with service issues and direct engagements between merchant and consumers.
- The Alime bot helps the end-use consumer. It is employed in online and phone hotline channels and relies on a rich set of interactive user interface components that can provide text dialogues, cards, graphics, videos, and other conversational interactions between robots and consumers.¹ Alime bot also possesses duplex voice dialogue

Table 1: Pros and Cons of Using AI or Human Agents for Customer Engagements

	AI	Human	Examples
Pros	<ul style="list-style-type: none"> • Faster and more accurate responses for cases with clearly defined questions and answers • Easily scalable, low marginal cost • High supply elasticity 	<ul style="list-style-type: none"> • More accurate responses for complex and non-standardized cases • Useful when customers need to detail evidence of problems • Better understanding of long-tail problems not well defined by algorithmic rules 	<ul style="list-style-type: none"> • Lead time: For Taobao's Double 11 shopping festival, human services typically need three months for hiring, preparation, and training in comparison to AI services that require only a one-day lead time focusing primarily on increasing cloud computing capacity. • Costs: AI-based customer service costs less than 1 percent of human customer service.
Cons	<ul style="list-style-type: none"> • Unable to handle complex scenarios with many rounds of conversation • Strong dependence on algorithms and training data • Weak emotional connection with customers 	<ul style="list-style-type: none"> • Slower responses • Quality is highly dependent on personnel • Low supply elasticity • High cost 	<ul style="list-style-type: none"> • Interpretability: AI-based bots need to be provided with at least hundreds of training samples; humans can read the question and (mostly) understand the problem • Number of interactions per customer: The current AI-based bot service is limited to two rounds of customer dialogue; humans can continue for longer customer dialogues

Figure 1. Alibaba's Five AI-bots Employed for Customer Service



capabilities and a targeted user interface to serve users who prefer to consult by telephone.

- The AI bot complements Alime by proactively engaging with end-user consumers and acting as an intermediary during service disputes with business merchants. The AI bot's underlying algorithms use transaction and conversational information to evaluate service disputes and make automated judgments. The AI bot then calls end-users to discuss the dispute decision. If consumers are satisfied with the AI bot's decision, the case is closed; if not, the AI bot helps consumers fill out appeal forms to have the case resolved by a human evaluator.
- The Dahuang bot helps train customer service personnel by simulating a larger and more diverse set of consumer and merchant encounters than a non-AI training system. (See Figure 1.)

Implementation challenges and solutions

The first challenge Alibaba faced was organizational hesitancy: Engineers and leaders were wary of AI's ability to positively engage with and resolve consumer and merchant queries. To overcome this hesitancy, the firm implemented a fast-fail strategy² that contin-

The first challenge Alibaba faced was organizational hesitancy: Engineers and leaders were wary of AI's ability to positively engage with and resolve consumer and merchant queries.

ually tested the AI against human responses in small real-world experiments.

The pilot proved that the AI chatbots outperformed their human counterparts, delivering superior customer satisfaction scores and improving merchants' first-contact resolution scores. Those results led the organization to fully buy in to implementing AI customer engagement services.

The second challenge involved two technical issues. Many initial customer queries involved a user-intent classification problem: Queries were stated similarly but had different underlying intents. For example, "I need help with my order" can mean needing help with tracking an order, getting a refund, or many other intents. In addition, many customer queries involved a long-tail problem: extremely low likelihood of certain types of queries. Since niche topics by nature gener-

ate less data, queries about them tend to lead to less accurate chatbot answers – which lowers customer satisfaction.

To overcome the first technical problem, Alibaba developed an extensible multitask learning paradigm using a meta lifelong learning framework that learned robust text representations across tasks and employed a least recently used (LRU) replacement policy to manage model deployment and memory resources.³ To overcome the latter problem, Alibaba implemented a multi-grained interactive matching network for few-shot text classification that leverages a dynamic routing algorithm in meta-learning to better adapt and generalize unseen classes while also providing more memory-based flexibility.⁴

AI vs. human results

Overall, Alibaba's AI-based customer satisfaction scores exceeded or matched human interactions in most product categories. Crucially, it provided customers much quicker responses and enabled Alibaba to engage with customers at all hours of the day.

A/B testing results reveal that, in its first few weeks, Alibaba's proactive AI-based intermediary dispute service resulted in a 25 percent increase in customer satisfaction over its previous non-AI-based dispute resolution procedures.

Meanwhile, Alibaba's training bot is assisting more than 1,500 customer service personnel daily. In addition, it is reducing customer service personnel training time by more than 20 percent.

Finally, Alibaba has realized cost savings of over a billion RMB annually (~US \$150 million) by employing AI over humans for service engagements.

Lessons learned and next steps

Alibaba employs its five AI-based chatbots in more than eighty

Table 2: Summary of Alibaba's Five AI Bots

AI-bot	Alibee Shop	Wanxiang-bot	Alime (bot & phone hotline)	Alime (customer service dispute)	Dahuang-bot
Target Customer	End-user consumers	Business merchants	End-user consumers	End-user consumers	Customer service personnel
Year Service Began	2018	2015	2015 (online) 2018 (hotline)	2021	2017
Purpose	Assist consumers with merchant-related questions on Taobao platform (via chat)	Assist merchants with help about the Taobao platform (via chat)	Assist end-user consumers use of the Taobao platform (via chat and voice hotline)	Assist end-user consumers service disputes with business merchants (via chat and voice hotline)	Training customer service personnel with a broader and more diverse range of questions than human-based training (via chat and voice hotline)
Examples	Service engagements involving direct merchant-to-consumer engagements	Service engagements involving business merchant questions on the platform's rules, activities, and service issues	Service engagements involving customer service questions and customer engagements	Service engagements resolving dispute resolutions	Training customer service personnel on how to respond to customer service engagements
Key Technical Features	<ul style="list-style-type: none"> • QA match • Intention recognition • Rec. system 	<ul style="list-style-type: none"> • QA match • Intention recognition 	<ul style="list-style-type: none"> • QA match • Intention recognition (bot) • Visual QA (bot) • Automated speech recognition (hotline) • Multi-turn dialogue (hotline) • Text-to-speech (hotline) 	<ul style="list-style-type: none"> • QA match • Intention recognition 	

Question-Answer (QA) match: the need to match the user's query to the title of the corresponding knowledge base. The user's description is based on natural language, which has rich and diverse expressions for the same problem. Only AI models based on deep learning can provide highly accurate and normalized solutions based on the quantity of knowledge data need to address the user's questions.

Multi-turn dialogue: the ability for an AI bot to engage in back-to-back conversational turns with a user.

Alibaba-related apps, including Taobao, Xianyu, Fliggy, Hema, and Lazada. It took Alibaba developers one year to create the initial AI bot. However, this process has become much more efficient, and it now takes Alibaba's developers only one month to develop a new AI bot for its platforms, although they have to continually improve each bot.

Alibaba has also learned that customers have accepted the chatbots and appreciate their role in resolving many, though not all, types of customer service queries. It has learned that it's critical to constantly test the structure to see what types of help customers are willing to accept from AI chatbots.

A key lesson Alibaba learned was that, despite its chatbots'

success, AI cannot and should not completely replace human customer service agents. But each should be deployed in the scenarios that best suit their abilities.

For simple FAQ questions, AI can directly reply to users. However, for complex complaints and disputes, AI can attend to labor-intensive tasks, such as collecting informa-

Alibaba has realized a cost savings of over a billion RMB annually (US \$155 million) by employing AI over humans for service engagements.

tion on appeals and vouchers, and possibly make initial decisions, but humans will still need to review the information, as well as related non-standardized materials, to make final judgments.

Going forward, Alibaba will continually invest significant resources in making a seamless human-machine collaboration. For example, Alibaba's AI bots continuously monitor whether customers encounter obstacles and whether the AI-based service can understand and resolve their queries.

Whenever appropriate, the AI bot automatically transfers customers to human service agents and then prompts and provides those agents with essential information to avoid asking the

customers to repeatedly describe the problem.

At the same time, the AI bot enhances customer interactions with human agents by providing data-informed solutions and particular phrasing recommendations, as well as issuing immediate warnings to human agents if they behave improperly.

Finally, Alibaba believes that, although technologically proficient AI natural language processing models and training data are important, establishing an organizational mind-set which offers customer-centric AI-based solutions to efficiently solve customer problems is more important than just implementing AI.

So, Alibaba is continuing to refine its AI bots through hundreds of releases, iterations, and improvements over time, with the ultimate goal of solving customers' problems in the most efficient and satisfying method possible, whether through AI or human service. ■

Author Bios



Yitong Wang is head of the Data Science and Business Insights Department in the Customer Experience Business Group at the Alibaba Group. Yitong earned his undergraduate and master's degrees from Tsinghua University and his PhD from the University of California, Irvine. His research has been published in the *Journal of Consumer Psychology*, *Strategic Management Journal*, and *Journal of Experimental Psychology: General*, among others.



Depin Chen is the Director of Applied Algorithms in the Customer Experience Business Group at the Alibaba Group. Depin is an expert in developing recommender systems, conversational AI, anomaly detection, and more. He earned his Bachelor's degree and PhD the University of Science and Technology of China, in 2005 and 2010, respectively. He has several publications in leading data and computer science conference proceedings such as ICDM, SIGIR, and PAKDD.



Ofer Mintz is an Associate Professor of Marketing at the University of Technology Sydney, Visiting Associate Professor of Marketing at Tel Aviv University, and author of *The Post-Pandemic Business Playbook: Customer-Centric Solutions to Help Your Firm Grow*. His research focuses on marketing metrics, analytics, strategies, and marketing's role in start-up firms. His research has been published in the *Journal of Marketing*, *Marketing Science*, and *World Economic Forum*, among other managerial and academic outlets.



Kehan Chen is Manager of the Chatbot Algorithm Team in the Customer Experience Business Group leading the AliMe Conversation AI algorithm team at the Alibaba Group. Kehan's research interests include natural language understanding, human-computer interaction, and dialogue systems. Kehan earned his master's degree from Zhejiang University, Hangzhou, China and has authored or coauthored several papers in leading data and computer science conference proceedings, including at Interspeech and KDD.

Endnotes

1. Alibaba's AI-bot's interactive UI components are open sourced and available for others to employ (see <https://chatui.io/>).
2. For further details on fast-fail strategies, please see: Mintz, Ofer, and Eric Knight (2022), "Beyond Usual: [Six] Ways Leading Firms Diverge from Business as Usual." *Management and Business Review*, forthcoming.
3. For further details, please see: Wang, Chengyu, Haojie Pan, Yuan Liu, Kehan Chen, Minghui Qiu, Wei Zhou, Jun Huang, Haiqing Chen, Wei Lin, and Deng Cai (2021), "MeLL: Large-scale Extensible User Intent Classification for Dialogue Systems with Meta Lifelong Learning," in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, New York, NY, USA: Association for Computing Machinery, 3649-3659.
4. For further details, please see: Geng, Ruiying, Binhua Li, Yongbin Li, Jian Sun, and Xiaodan Zhu (2020), "Dynamic Memory Induction Networks for Few-Shot Text Classification," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online: Association for Computational Linguistics, 1087-1094.