



Beyond Conversational Artificial Intelligence

Mark Campbell, EVOTEK

Humans have struggled for centuries to give voice to machines. Since conversational artificial intelligence (AI) emerged, it has changed the business models in many industries, and this is only the beginning. How did conversational AI develop, how does it work, and where is it going?

Since machines were invented, humankind has strained to hear the vox machina, or the voice in the machine. During the 1200s, medieval Kabbalists tried using combinatorics on the Hebrew alphabet to discover secret knowledge. During the 1700s, Leibniz abandoned his quest for a system of language and cognition, while Swift satirized the idea with “the machine” in *Gulliver’s Travels*.¹ As the digital age dawned during the 1950s, Alan Turing devised the “imitation game” between humans and machines (that is, the Turing test) to determine “if machines think.”²

Joseph Weizenbaum’s 1966 program Eliza made “natural language conversation with a computer possible.”

Named after the protagonist in Shaw’s *Pygmalion*, the system followed the interrogational style of a Rogerian psychotherapist, asking open-ended questions sculpted around keywords in the user input. Despite its simplicity, Eliza made quite an emotional connection with many users who poured out their innermost issues to an infinitely patient and empathetic counselor [for as long as the timeshare on the Massachusetts Institute of Technology (MIT) mainframe allowed, anyway].³

Eliza was followed by a cadre of ever-increasingly sophisticated “chatbots,” such as PARRY, Jabberwacky, Dr. Sbaitso, ALICE, SmarterChild, and, finally, Siri, the first widely adopted virtual assistant, in 2010.⁴ With the advent of more humanlike linguistic artificial intelligence (AI) techniques, large online companies launched virtual assistants, including Alexa, Cortana, Messenger, and Tay (who quickly developed into a foul-mouthed, bigoted misogynist until her plug was pulled).⁵

THE AGE OF CONVERSATIONAL AI

Unfortunately, digital users seldom converse with convenient keywords in proper, grammatically correct sentences. Instead they often misspell, use slang, make typos, ask compound questions, and assume an implied context from previous statements. Thus, modern-day Elizas have

had to go far beyond the decision tree responses and scripted dialogue of the earlier chatbot era—they must be conversational. Today, human-machine interaction is built on, or enhanced by, a variety of AI approaches, especially

and language processing courses are rewritten each year to reflect the discoveries of their graduates.⁶ Nonetheless, many conversational AI solutions are built around four common components (illustrated in Figure 1).^{8, 9, 10}

During the 1200s, medieval Kabbalists tried using combinatorics on the Hebrew alphabet to discover secret knowledge.

deep learning methods, hence the term “conversational AI.”⁶ Solutions employing AI have outstripped classic linguistic and rules-based approaches, and those trained on large proprietary data sets have progressed the fastest. Trained on more than 300 Gb of social media texts, Google’s Meena can conduct discussions on any open-ended topic—and even throw in a (bad) pun.⁷

It is an understatement to say that conversational AI is evolving rapidly. At Stanford and MIT, many speech

- **Request:** This ingests speech from the user and typically incorporates automatic speech recognition to convert speech to text, and it includes a traffic manager to route conversation text to the correct processing instances.
- **Natural language understanding (NLU):** The NLU component performs tasks to make it possible to construct a coherent response to the user.

- **Natural language processing (NLP):** Depending on the complexity of the response and the required accuracy of the system, the NLU may need to parse various aspects of the user’s request through an NLP unit. Although NLPs are typically proprietary and vary considerably, most take user input through a semantic grinder to extract key components. For instance, the three inputs of “I want to go to Cleveland,” “I need to buy a plane ticket,” and “Oh, I’ll need to book a hotel,” could be decomposed as follows:

- **Domain:** pertains to the classification of task groups to which the user is referring (for example, *book hotel* and *buy plane ticket*)
- **Intent:** relates to one of the domain’s supported actions (for example, *buy plane ticket*)
- **Entity:** relates to one of the domain’s supported objects (for instance, *buy plane ticket*)
- **Role:** a secondary entity classification to distinguish between similar elements (for example, *buy a one-way plane ticket* versus *buy a round-trip ticket*)
- **Context:** intent and entity correlation from past input (for instance, “I want to go to Cleveland” changes the entity to *buy a plane ticket to Cleveland*)
- **Grouping:** a language parser to determine related entity groups (for instance, *book hotel in Cleveland* and *buy a round-trip plane ticket to Cleveland*).

- **Dialogue manager and interpreter:** This examines the dialogue flow to gather context, elicit user feedback, and learn what actions most effectively

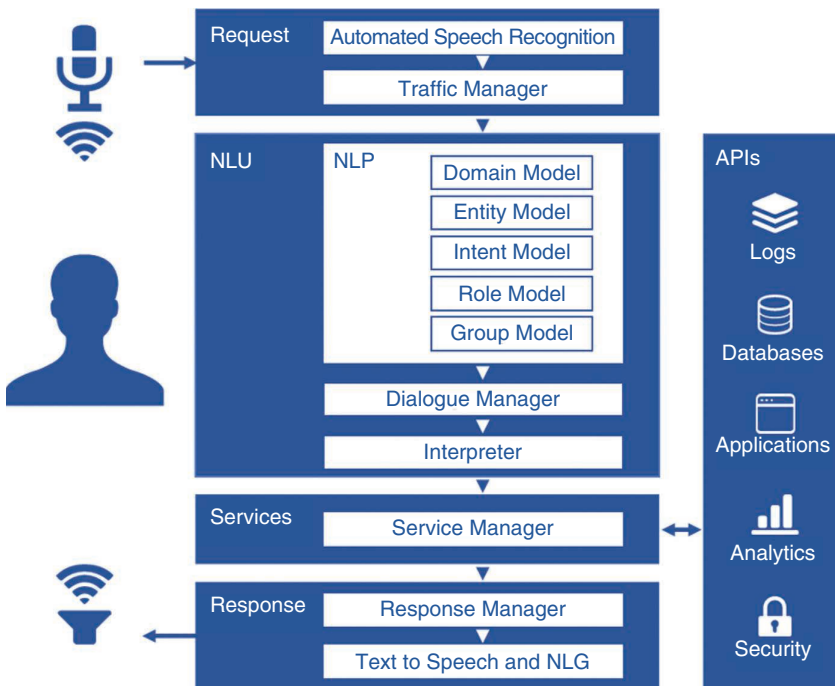


FIGURE 1. The conversational AI architecture. NLU: natural language understanding; NLP: natural language processing; API: application programming interface; NLG: natural language generation.

satisfy user requests. Sentiment analysis is another prevalent feature of dialogue management that monitors the user's mood and emotional state. A rapid escalation in user emotions can cause the NLP to hand the conversation over to a human operator to calm the situation. Conversely, positive emotional interactions are fed forward to the response manager to reinforce "happy path" interactions.

- › **Services:** The goal of all conversational AI is to transform user conversation into business transactions. To do this, a service manager integrates with application programming interfaces for back-end business applications, databases, security policies, audit logs, analytics, and visualization systems.
- › **Response:** This generates and provides responses to the user and commonly contains the following:
 - **Response manager:** This takes the conversational components and generates response candidates from which to select the best answer to send to the user. Two response manager models are prevalent:
 - **Retrieval-based models:** As today's most common response model, these select candidate answers from a knowledge base of previously observed patterns in intent, entities, and context. For example, "Which of these flights to Cleveland would you prefer?"
 - **Generative models:** Growing in popularity, these create response candidates from user input and previous conversations using a natural language generation (NLG) model. NLG models are rapidly becoming more sophisticated and now outperform

retrieval models, as they can create proactive suggestions based on past transactional patterns. For example, "Hey, the Browns have a playoff game while you're in Cleveland. Do you want tickets to the game?"

- **Text to speech and NLG:** A text-to-speech converter produces the "voice" response to

method of last resort. Conversational AI, such as Gridspace's Grace, has flipped this model upside down and can now provide one of the fastest, cheapest, and most accurate routes to resolving customer issues.⁶

Conversational AI is entering many markets that otherwise could not technically or financially afford to integrate direct human-machine integration into their business transactions.

Eliza made quite an emotional connection
with many users who poured out their
innermost issues to an infinitely patient and
empathetic counselor.

the user based on candidates selected by the response manager. Depending on the level of complexity of the response to the user, the text-to-speech converter may be augmented with an NLG module.

CONVERSATIONAL AI IN TRANSITION

It's an exciting time for conversational AI, as armies of cyberagents, virtual assistants, and digital advisers are rapidly eradicating traditional chatbots. The conversational AI market is expected to grow from US\$4.8 billion in 2020 to US\$13.9 billion in 2025 at an annual growth rate of 21.9%.¹¹ Outside the familiar domains of home assistants and smartphone personas, conversational AI pervades contact centers, help desks, conversational commerce, and application interfaces.

The conversational AI transition is changing the underlying business models in many industries. Call center operations, for example, have traditionally been the most costly, unsatisfying, and time-consuming customer support method. Thus, companies have historically worked to redirect customers to e-mail, FAQ databases, and annoying interactive voice response option trees—a human agent was the

It's been rapidly adopted in interactive customer support, sales and marketing, and conversational commerce for local merchants. One credit card swipe gives small retailers a customer interaction structure that rivals or surpasses those found in large international powerhouses. Many interactive applications employ conversational help features that examine what the user has done (context) and what he or she is trying to do (intent) and verbally guide him or her to completion.

As conversational AI expands into areas such as employee onboarding, complex commerce, direct marketing campaigns, and mobile access, its internal mechanisms are also in transition. As mentioned, response management is moving from response retrieval to generative response models or a combination of both. This subtle change enables conversational AI to mature from a reactive servant of user requests to a proactive and intuitive participant in user dialogue. For example, using only conversational AI and a browser, Ushur's Invisible App guides users through complex business processes without the need for development, testing, and downloads.

Conversational AI is also flipping the request-response cycle so that back-end systems can initiate the

request (for example: rates are low, do you want to go to Cleveland?), push it to the user, and await a response.¹² Intent parsing is also maturing from single-dimensional sentence dissection to multilevel categorization. Multilevel

communication between a pilot and a controller. Regulatory hurdles aside, AI-controlled air traffic control conversations will be calmer, more accurate, and, ultimately, much safer

evolve for a long time to come. Direct neural connections, anyone? **■**

It's an exciting time for conversational AI, as armies of cyberagents, virtual assistants, and digital advisers are rapidly eradicating traditional chatbots.

intent identifies the purposes and conversational context of compound requests (referred to as the *stemmed form*) and uses a scoring algorithm to predict the predominant objective, asks the user for any additional information, completes the primary objective, and repeats the process for the next-strongest goal.¹³

BEYOND CONVERSATIONAL AI

Conversational AI will undoubtedly continue permeating all business activities that require human interaction, and it has great potential in the following:

- **Co-automation:** Conversational AI will act as the link between humans and automation systems in collaborative processes that take advantage of what people and machines each do best.¹⁴
- **The Internet of Things:** As conversational AI systems are optimized, tuned, and miniaturized, they will be embedded into many smart edge devices, such as vehicles, appliances, facilities, security equipment, and sensors, giving those mechanisms the ability to converse directly with operators.
- **Air traffic control:** Before you freak out, know that conversational AI is a natural fit for air traffic control, where the interaction is verbal

than the current human-human system.¹⁵

- **Geriatric care:** Just imagine the myriad uses of a conversational companion for the growing elderly population: an ever-patient, smart assistant conversing on a variety of topics, reminding people about appointments and medications, alerting first responders in case of an emergency and sharing pertinent health-care information, and virtually reminiscing about a meaningful song or photo album.
- **Digital humans:** Companies such as Uneeq are coupling conversational AI with animated AI-generated faces to augment the virtual assistant persona. By combining facial gestures and body language with conversation, these digital humans enhance the conversational experience while increasing communication accuracy and customer satisfaction.¹⁶

As with any AI technology, conversational AI awaits the inevitable challenges of security, privacy, unintentional bias, and ethical gray areas. But the sweeping adoption of conversational AI we are witnessing today, coupled with its incredible potential uses, ensures that Eliza's descendants will thrive and

REFERENCES

1. O. Schwartz, "In the 17th century, Leibniz dreamed of a machine that could calculate ideas," *IEEE Spectrum*, Nov. 4, 2019. [Online]. Available: <https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/in-the-17th-century-leibniz-dreamed-of-a-machine-that-could-calculate-ideas>
2. A. Turing, "Computing machinery and intelligence," *Mind*, vol. LIX, no. 236, pp. 433–460, 1950. doi: 10.1093/mind/LIX.236.433.
3. J. Weizenbaum, "Computational linguistics," *Commun. ACM*, vol. 9, no. 1, pp. 36–45, 1966. doi: 10.1145/365153.365168.
4. "The history of chatbots." *Futurism*. Accessed: Sept. 3, 2020. [Online]. Available: <https://futurism.com/images/the-history-of-chatbots-infographic>
5. S. Perez, "Microsoft silences its new A.I. bot Tay, after Twitter users teach it racism," *TechCrunch*, Bay Area, Mar. 24, 2016. [Online]. Available: <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/>
6. M. Campbell, interview with E. Macmillan, Aug. 27, 2020.
7. E. H. Schwartz, "Google's new Meena chatbot imitates human conversation and bad jokes." *Voicebot*. [Online]. Available: <https://voicebot.ai/2020/02/03/googles-new-meena-chatbot-imitates-human-conversation-and-bad-jokes/>
8. A. Smith, "Understanding architecture models of chatbot and response generation mechanisms." *DZone*. Accessed: Sept. 3, 2020. [Online]. Available: <https://dzone.com/articles/understanding-architecture-models-of-chatbot-and-r>
9. MindMeld, "The conversational playbook: Platform architecture." Accessed: Sept. 4, 2020. [Online].

- Available: <https://www.mindmeld.com/docs/userguide/architecture.html#:~:text=The%20MindMeld%20Conversational%20AI%20Platform,intelligent%20data%2Ddriven%20conversational%20apps.&text=The%20Application%20Manager%2C%20while%20part,and%20never%20needs%20developer%20attention>
10. "How do chatbots work? A guide to the chatbot architecture." Marutitech. Accessed: Sept. 3, 2020. [Online]. Available: <https://marutitech.com/chatbots-work-guide-chatbot-architecture/>
 11. "Conversational AI market by component (platform and services), Type (IVA and chatbots), technology (ML and Deep Learning, NLP, and ASR), application, deployment mode (cloud and on-premises), vertical, and region: Global forecast to 2025." Researchandmarkets. [Online]. Available: <https://www.researchandmarkets.com/reports/5136158/conversational-ai-market-by-component-platform>
 12. M. Campbell, interview with M. Fisher and H. Peter, Sept. 2, 2020.
 13. "Using conversational AI to decode complex customer interactions." Boost. Accessed: Sept. 4, 2020. [Online]. Available: <https://www.boost.ai/articles/using-conversational-ai-to-decode-complex-customer-interactions>
 14. L. James, "Why humans and machines work better together," IT Pro. [Online]. Available: <https://www.itpro.co.uk/technology/artificial-intelligence-ai/355573/why-humans-and-machines-work-better-together>
 15. P. Dialani, "Leading-edge applications of AI and machine learning." Analyticsinsights. [Online]. Available: <https://www.analyticsinsight.net/leading-edge-applications-of-machine-learning-and-artificial-intelligence/>
 16. "UneeQ drives new human to machine interface with unveiling of first-of-its-kind digital human creator platform." BusinessWire. <https://www.businesswire.com/news/home/20200728005336/en/UneeQ-Drives-New-Human-Machine-Interface-Unveiling>

MARK CAMPBELL is the chief innovation officer at EVOTEK. Contact him at mark@evotek.com.



IEEE Security & Privacy magazine provides articles with both a practical and research bent by the top thinkers in the field.

- stay current on the latest security tools and theories and gain invaluable practical and research knowledge,
- learn more about the latest techniques and cutting-edge technology, and
- discover case studies, tutorials, columns, and in-depth interviews and podcasts for the information security industry.



computer.org/security

Digital Object Identifier 10.1109/MC.2020.3034003