Conversational AI is reshaping the human-machine interaction

November 2020
What is Conversational AI?
Conversational AI (CAI) combines natural language processing, AI, and machine learning to understand and respond to free-form text or voice in an engaging and personalized manner.

**Consumer application**
- **Smart phone**
  - Voice assistant APP
- **Smart home**
  - Intelligent lighting
  - Intelligent kitchen
- **Wearable device**
  - Smart watch
  - Smart band
- **Intelligent vehicle**
  - Speech navigation
  - Voice control

**Business application**
- **Healthcare**
  - Voice diagnosis
  - Hospital guidance
- **Education**
  - Speech teaching
  - Speech assessment
- **Finance**
  - Call center
  - Interactive voice response
- **Manufacturing**
  - Industrial robot

**Typical scenarios**

Transform the ways we live and work
Benefits of CAI in organizations

• Generate quick responses to reduce waiting time and processing time
• Reduce operational costs such as labor cost and time cost
• Answer questions and process tasks in a uniform format
• Reduce human errors with more consistent services

Now

Standardization

Efficiency

Experience

Diversity

Future

Personalization

• More diverse interaction methods can be achieved, such as text, voice, gesture, etc.
• CAI can be applied to scenarios where the traditional interaction methods could be inconvenient, such as self-driving
• CAI will change ways how human and machine interact, optimize task-based working processes, and can tremendously reduce the time for information retrieval
• Being a productive personal assistant, CAI will augment human capabilities in the future of work

• Use CAI to replace manual operations enabling employees to focus on more creative tasks
• Interact with customers smoothly and efficiently creating better customer experience
• More diverse interaction methods can be achieved, such as text, voice, gesture, etc.
• CAI can be applied to scenarios where the traditional interaction methods could be inconvenient, such as self-driving
• CAI will change ways how human and machine interact, optimize task-based working processes, and can tremendously reduce the time for information retrieval
• Being a productive personal assistant, CAI will augment human capabilities in the future of work
CAI development
CAI growth in applications

CAI plays a major role in the applications of Artificial Intelligence

Leading use cases for CAI in AI deployment

- Facial recognition: 45%
- Call-Center virtual customer agents: 42%
- Chatbots: 39%
- Fraud analysis on transactional data: 37%
- Process optimization: 37%
- Market/Consumer segmentation: 37%
- Virtual personal assistants: 17%
- Smart robotics: 17%

(Percentage of China respondents)

CAI penetration in leading AI industries

- Finance: High penetration
- Retail: High penetration
- Self-driving: High penetration
- Communications: High penetration
- Government: High penetration
- Education: High penetration
- Healthcare: High penetration
- Smart City: High penetration
- Retail: High penetration

(Penetration: industry application degree; Market size: marketing opportunity)

- CAI has a higher penetration rate in finance, education, government and healthcare among applications across industries.

2. Global AI Development Whitepaper, Deloitte, September 2019

Speech-activated applications have been widely adopted in the field of AI, such as call-center virtual customer agents, chatbots, virtual personal assistants, and smart robotics.
Four drivers of CAI growth

1. Machine learning, deep learning and other technologies build solid foundation for CAI
   - Deep learning has made great breakthroughs in speech recognition, natural language processing and speech synthesis.
   - In the future, it is possible to realize barrier-free human-machine emotional communications.

2. Development of chips and cloud technology has fueled the basic computing power to CAI
   - The development of chips and cloud computing shows the trend of integrating with AI.
   - With the development of edge AI chips, CAI will find its way into mobile devices.
   - Cloud computing enables enterprises and governments to offer more personalized and intelligent services and products.

3. Policy is a catalyst for the future of AI growth
   - Policy has gone through three phases, from single products to data-driven innovation platforms, from individual actions to national strategies, and most importantly, from AI technology development to the integration of AI and the real economy.

4. CAI is transforming the ways we live and work
   - Extensive applications of CAI have emerged in consumer market.
   - CAI brings value enhancement to various industries.
   - Users hold more positive attitudes toward CAI applications.
Prior to 2011, when deep learning, big data and cloud computing were not integrated with speech recognition technology, the accuracy of speech recognition was 54.61%.

In 2011, Microsoft introduced deep learning into speech recognition, which improved the accuracy of speech recognition to 81.55%.

In 2017, the accuracy of Microsoft’s switchboard reached 94.99%, surpassing human for the first time.

Before 2011, voice applications developed by Microsoft and Google were only based on basic grammar analysis and machine translation, mainly in information retrieval and extraction.

Before artificial neural network had been used in speech synthesis technology, voice assistants started to imitate human intonation.

Before 2011, first-generation voice assistants were mainly used for navigating on PCs and information retrieval in a monotone voice.

In 2017, Tacotron 2, a speech synthesis system released by Google, was developed as close as human voice and became a benchmark system.

With the application of pre-training language model in NLP task, NLP technology started to focus on emotional text analysis and text reasoning.

After 2011, NLP text analysis was advanced towards deep understanding, making dialogue robots more practical and scene oriented.
In 2006, Amazon created AWS, providing a variety of cloud-based services on IaaS including storage and computation. Before 2007 when AI was in the early stage, CPU chips were sufficient to provide enough computing power.

After 2013, GPU were widely used for AI. In 2015, Google first released the ASIC chip tpu1.0, and the industry started to develop special chips for AI. FPGA came out after Intel acquired Altera in 2015.

In 2017, Huawei Qilin 970 became the first mobile phone AI chip, introducing AI into mobile devices. In 2016, AWS officially launched its own AI product line, and cloud computing started to show the trend of integration with AI.

Salesforce pioneered the enterprise PaaS market when it launched the Force platform in 2008. After 2010, SAP, Oracle and other traditional software companies began to launch cloud services, and enterprise SaaS developed rapidly.

The increasing computing power releases the potential of AI algorithms.

The continuous chip evolution integrated with cloud computing.

- Bringing AI to the device: Edge AI chips come into its own.
- Edge AI chips will find the way into consumer devices and enterprise markets, such as smartphones, smart speakers, wearables and robots, cameras, sensors, and other IoT devices in general.
- Services become more intelligent: virtual computing platform greatly improves the data processing and reduces the cost of using data.
- Enter the era of customization: based on the extensive collection of information, user behavior and needs are accurately pinpointed.
### Policy

**2015.05**  
Made in China 2025  
Develop intelligent equipments and products, and promote intelligent production processes

**2016.07**  
The 13th five-year plan for national technology innovation  
Focus on the development of AI technology driven by big data

**2016.09**  
Special action for innovation and development of intelligent hardware industry  
Focus on developing intelligent equipments and service robots

**2016.05**  
“Internet+” AI three-year action plan  
By 2018, create basic artificial intelligence resources and innovation platforms

**2017.03**  
Government work report  
Mention AI for the first time

**2017.07**  
New generation of artificial intelligence development plan  
AI become the national strategy

**2017.10**  
The report of the 19th national congress  
Promote the close integration of Internet, big data, AI and the real economy

**2017.12**  
Three-year action plan to promote the development of a new generation of Artificial Intelligence industry (2018-2020)  
Push further integration of AI with the real economy

**2018.09**  
List of innovation projects of deep integration of AI and the real economy  
Promote the integration of AI and the real economy

**2019.03**  
Guiding opinions on promoting the deep integration of AI and real economy  
Establish data-driven smart economy and boost innovation vitality

---

**Single product**  
Innovation platform  
Individual action  
National strategy  
Technology development  
Real economy integration
User demand - consumer

Why Conversational AI in consumer market?
- Hands-free
- Improve the quality of life
- Enhance human-machine interaction

Users hold a more positive attitude toward CAI

<table>
<thead>
<tr>
<th>Equipment</th>
<th>CAI Expectation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mobile phone</td>
<td>81.6%</td>
</tr>
<tr>
<td>Air conditioner</td>
<td>56.8%</td>
</tr>
<tr>
<td>Television</td>
<td>73.6%</td>
</tr>
<tr>
<td>Refrigerator</td>
<td>47.8%</td>
</tr>
<tr>
<td>Car</td>
<td>65.1%</td>
</tr>
<tr>
<td>Watch</td>
<td>40.7%</td>
</tr>
<tr>
<td>Speaker</td>
<td>25.2%</td>
</tr>
<tr>
<td>Robot</td>
<td>58.0%</td>
</tr>
</tbody>
</table>

User’s expectation of CAI for equipment

On smartphones, CAI related applications that users expect to see are information inquiry, weather broadcast, function setting and chatting.
On televisions, voice control switch and timer setting are the top two desired CAI applications.
In terms of intelligent vehicles, applications with conversational features expected by users are mainly voice navigation and control system.

Source: Conversational AI Whitepaper, IDC, 2017
Why Conversational AI in business market?

- Improve operational efficiency
- Reduce labor cost
- Enhance quality of services
- Transform the ways we work

Top AI benefits selected by organizations across the globe

- Making process more efficient: free up hands, change working paradigms and replace the tedious and high-cost operations with machines.
- Improving decision-making: find hot spots for the market and provide data support for follow-up planning.
- Enhancing existing products and services: transform human-machine interaction to provide customers with more personalized experience.

Source: Thriving in the era of pervasive AI, Deloitte, July 2020
Industry specific commercial applications
## Applications of CAI in Financial Services

### Mostly adopted applications of CAI in financial service

1. **Customer service**
   - **Online virtual assistants**
     Online virtual assistants play a complementary role to human agents. These virtual assistants would chat with customers, understand customers’ requirements and respond to their inquiries in an interactive way.
   - **Outbound call system**
     Outbound virtual agents would take the initiative to call customers, screening out most of the repetitive work for human assistant, such as tele-sales, customer feedback collection, and identity authentication.

2. **Speech navigation**
   - **Speech navigation**
     This application replaces manual press-button function with navigation via speech, making it easier for customers to navigate on the phone when calling to service center.

3. **Extensive analysis**
   - **Find hot spots in the market**
     With massive amount of speech information, identify most frequently used words and extract information valuable to potential commercial opportunities.
   - **Provide data support for machine learning**
     The corpus, generated from the speech information, greatly improves the machine learning database. Additionally, such database would provide reference for subsequent conversational interaction.

### Application scenarios embedded CAI in different fields of financial industry

#### Banking
- Provide counter services inquiries such as deposit, transfer, loan and account settings

#### Insurance
- Recommend appropriate insurance products according to user profile and experience

#### Investment
- Provide the latest price information of stocks and funds, and speculate about the future

### Industry specific commercial application – financial services

- In the financial industry, hundreds of banks, security companies and insurance companies have launched intelligent customer service systems since 2017.
- Due to huge total investment and fund size, Robo-advisor can be expected.
Industry specific commercial application – healthcare

CAI has penetrated in every process of the hospital care system

### Diagnosis

**Formation of electronic medical record**

Record the verbal consultation between doctors and patients, and generate the electronic medical records with CAI system processing voice information automatically.

Testing outcomes and radiology diagnosis reports can also be generated by physicians with voice narration.

### Operation

**Non-contact intelligent voice interaction**

In the operating room, doctors can retrieve the medical records, images and other relevant data of patients through speech input.

**Medical robot**

Compare the patient’s disease description with the standard medical guide, and provide services such as guidance, medical consulting, self diagnosis, etc. At the same time, the intelligent robot can complete medical history collection work in advance through dialogue.

### Reporting

**Speech generation of diagnosis report**

Integrate with each diagnosis report workstation to realize speech generation of diagnosis report, and transmit diagnosis results to patients and doctors in the form of voice output.

CAI applications have improved the healthcare services and alleviated the medical pressure caused by the shortage of medical resources with more efficient solutions.
Industry specific commercial application – automobile

**Voice will become the primary interaction media between human and auto**

**Voice navigation**
The system can understand the travel requirements of customers through voice interaction and complete the navigation service, which helps the driver focus on driving and ensures the safety to the maximum extent.

**Control system**
Instead of pressing buttons to control the car, drivers and passengers can directly control the car by talking to it, enjoying a more comfortable and convenient ride, improving the safety during driving and enriching the entertaining experience.

**Search engine**
Replace the traditional search interface with voice interaction, including music search, weather and calendar query, stock search, restaurant recommendation and other functions. It makes the vehicle more intelligent and improves the in-car experience.

**Security**
Based on the voice recognition and monitoring system, the on-board equipment can send out security alarm to the emergency contacts, so as to maintain the safety of vehicles and passengers and improve the lack of safety services in the vehicle system.

In 2019, the penetration rate of CAI in intelligent vehicles has reached 40%, and shows a continuous growth trend.

Source: Conversational AI Whitepaper, IDC, 2017
## Industry specific commercial application – education

CAI enhanced education will be adaptive, customized and unlimited

<table>
<thead>
<tr>
<th>Language</th>
<th>Practicing</th>
<th>Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAI can provide qualified and customized courses</strong></td>
<td>Pronunciation, passage reading, oral expression and etc.</td>
<td>Test and evaluate users' speech expression</td>
</tr>
<tr>
<td>Voice enhanced AI could replace traditional manual teaching in fields of both language and specialty education by providing qualified courses.</td>
<td>Users can practice oral pronunciation and Q&amp;A in multiple languages with CAI education system which will provide assessment, feedback and correction.</td>
<td>Assess the examinee’s pronunciation and language ability without human bias.</td>
</tr>
<tr>
<td>Intelligent adaptive course systems can use big data and algorithms to develop a set of effective and customized courses for all levels.</td>
<td>Structured musical practice</td>
<td>Proficiency and accuracy</td>
</tr>
<tr>
<td>By voice interaction and motion capture, students’ proficiency and accuracy in vocal, musical instruments and other skills are evaluated.</td>
<td>Based on the speech recognition and motion capture technologies, the comprehensive ability of examinees is evaluated, including proficiency, accuracy and etc.</td>
<td></td>
</tr>
</tbody>
</table>

- **Alleviate the problem of uneven distribution of education resources**: Intelligent education makes it possible to share educational resources.
- **Promote customized learning applications**: User demands drives educational products to be more diversified and personalized. The combination of online and offline courses makes it more adaptive to users at any level.
### CAI business applications landscape

<table>
<thead>
<tr>
<th>Industries</th>
<th>Value</th>
<th>Typical applications</th>
<th>Application cases</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Finance</strong></td>
<td>• Improving internal operations</td>
<td>• Call center</td>
<td>• Chat-bots replace agents and reduce their repetitive work</td>
</tr>
<tr>
<td></td>
<td>• Improving products and services</td>
<td>• Intelligent customer service</td>
<td>• Virtual assistants recommend personalized products</td>
</tr>
<tr>
<td></td>
<td>• Enhancing existing products</td>
<td>• Intelligent office system</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Making processes more efficient</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Optimizing decision-making</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Enhancing relationship with customers</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Enabling new business models</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Lowering costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Making employees more productive</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Healthcare</strong></td>
<td></td>
<td>• Electronic medical record</td>
<td>• Record the consultation processes and generate the electronic medical records</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mobile medical care</td>
<td>• Nurses can extract the patient’s information by voice input</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Diagnosis report</td>
<td>• Transmit diagnosis to patients and doctors in the form of voice</td>
</tr>
<tr>
<td><strong>Retailing</strong></td>
<td></td>
<td>• Intelligent logistics</td>
<td>• Real-time tracking of transportation path of goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intelligent sales</td>
<td>• Chat robot recommends personalized products to consumer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intelligent customer service</td>
<td>• Chatbot replaces manual customer service</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intelligent interface</td>
<td>• Interact with users through voice, video and other modes</td>
</tr>
<tr>
<td><strong>Manufacturing</strong></td>
<td></td>
<td>• Speech teaching</td>
<td>• Online intelligent teaching</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Speech evaluation</td>
<td>• Test and assess in the form of voice</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td>• Intelligent government service</td>
<td>• Optimize the public hotline</td>
</tr>
<tr>
<td><strong>Government</strong></td>
<td></td>
<td>• Integrate government service</td>
<td></td>
</tr>
</tbody>
</table>
Customer experience
CAI application status in customer service

Virtual Customer Assistant (VCA) market

VCA market growth
The Virtual Customer Assistant market will grow between 25% and 45% annually over the next three years.

Customer service interaction
By 2021, nearly one in six customer service interactions globally will be handled by AI.

Efficiency improvement
By 2025, customer service organizations that embed AI in their multichannel customer engagement platform will elevate operational efficiency by 25%.

Pioneering industries deploying VCAs

- Financial Services
- Technology, Media & Telecommunications
- Health Care
- Energy, Resources & Industrials

The mutual characteristics of those industries

- A large customer base frequently engaged
- A limited set of less-complex frequently asked questions fit for robots
- The continuous pursuit for reducing cost

Source: Gartner, Forrester, 2020
CAI challenges in customer service

What challenges will we encounter in customer services?

Lack of synergy between different channels
Fail to deliver better customer experiences
Lack of standardization of operations process
Limited functionality
Lack of human and AI blending
Long development period

Hard to deliver
Labor cost is still high
Employees may worry about their job prospects
Customers are reluctant to use chatbots instead of agents when seeking service
Cannot satisfy users' diverse demands
Efficiency has not been improved
Pain points of AI management in customer service

**Strategy and Planning**

Most enterprises do not have a holistic **data science strategy** to systematically plan and organize datasets, models and applications.

Customers leave information in multiple channels such as text, voice interaction and operation records. However, there is no **integration to deliver synergy**.

Most enterprises highly rely on vendors’ capabilities instead of **accumulating core AI competence** such as AI specialists, datasets, algorithm and etc.

**Procurement and Implementation**

Lack of appropriate **procurement methodology** in selecting vendors in data science.

The degree of **vertical specialization** is low in most standard VCA products, which is hard to meet user requirements in different businesses.

Lack of **data science middleware** to centralize data, prepare datasets, and train and manage models.

**Operation**

VCA **operation process** is not well designed. VCA only participates in basic work with high repetition. It still needs agents to participate in complex issues.

Lack of holistic **analytic functions** to monitor VCA performance.

Short of **maintenance and optimization of algorithms and models** after the deployment of VCA.
Data science middle platform

**Big Data Application**
- Visual reports
- Performance management
- Precision marketing
- Business intelligence
- Intelligent risk mgt.
- Intelligent logistics
- Intelligent customer services
- Personal assistant
- Robo-advisor

**Application**
- Capacity
  - Build corporate data science capacity
- Empower corporate business lines

- Efficiency
  - Code reuse to improve application development efficiency and reduce trial and error cost
  - Data integration to improve operation efficiency

- Security
  - Data asset protection
  - Authentication and access management
  - Traceable data link control

**Middle Platform**
- Data Processing
  - Data preparation
    - Data combination
    - Data quality
    - Data partitioning

  - Advanced data preparation
    - Data boxing
    - Data smoothing
    - Filtering and searching
    - Aggregate transformation

- Data Analysis
  - Interactive data exploration
    - Significance test
    - Clustering by self-organizing map
    - Geolocation image mapping
    - Correlation analysis
    - Investigative analysis
    - Probability density estimation
    - Similarity measurement

  - Interactive visual data analysis
    - Visual data exploration
    - Chart automation
    - Interactive dashboard
    - Customized optimization
    - Monovariant analysis

- Model Training
  - Basic machine learning
    - Regression
    - Classification
    - Clustering
    - Time-series analysis
    - Recommendation systems
    - Bayesian analysis
    - Composing methods

  - Advanced machine learning
    - Probabilistic graphical model
    - Knowledge graph
    - Deep learning
    - Transfer learning
    - Reinforcement learning

  - Feature engineering
    - Feature extraction
    - Feature transformation
    - Feature selection

  - Auto-ML
    - Business process automation
    - Hyperparameter optimization
    - Automated feature engineering
    - Neural architecture search

**Model Services**
- Model generation
  - Model warehouse
  - Model evaluation
  - Operation and maintenance

**Data Lake**
- Internal
  - OA
  - ERP
  - CRM
  - IoT
  - HR
  - Social media
  - Market monitoring

- External

**Computing Resource**
- Storage management
- Hardware
- Virtual machine
- Network resources
Process design: new human-machine cooperation model

### Replace

- **Customer asks question**
  - Virtual Assistant will replace the agent to help customers to solve problems.
- **Problems not solved**
  - Problem solved
  - Switch to the agent

- **This process will use Virtual Assistant instead of the agent to interact with customers directly to understand their needs and integrate relevant information to solve problems. Customer experiences will be directly affected by the voice technology.**
- **When Virtual Assistant and the agent have not reached a good connection and the problem cannot be solved, customers need to switch to the agent and repeat their needs. This process will not only decrease the efficiency of customer service, but also cause the negative customer experience.**

### Augment

- **Customer asks question**
  - The agent can push suggestions to the customer, or rejects suggestions and types their own answer.
  - Virtual assistant will give the solution that based on extracted information and data to the agent.

- **This process changes the traditional form of using Virtual Assistant to replace the agent. Virtual Assistant will support the agent without facing customers directly.**
- **Virtual Assistant will generate the answer automatically by analyzing customers’ needs and extracting information. Virtual Assistant can provide solutions for reference to the agent so as to increase their working efficiency.**
- **The agent can adopt the answer given by Virtual Assistant or reply independently. The solution will be more flexible and improve the customer experience.**

### Intermingle

- **Virtual Assistant will confirm customers’ basic information**
  - Virtual Assistant will evaluate customers’ needs and extract relevant information.
- **Customer asks question**
  - Virtual Assistant will send the previous record to the agent.
  - The agent will help customers to solve problems.

- **This process will use Virtual Assistant instead of the agent to finish basic tasks with high repetition rates, which leads to the good combination between Virtual Assistant and agent.**
- **Parts of labor can be liberated by this process and the efficiency of solving problems will be increased.**
- **Because of the accurate human-machine interaction, the agent can communicate with customers directly by using the interaction records, which will make the rapid improvement of customers’ experience.**
Experience optimization

- The **confidence level** (relevance) of related answers for every question is recorded automatically.
- For the questions with **low** confidence level answers, machines can send them to agents.
- **Manually optimize** the answer contents and send them back to machine for further training.

**User feedback collection**
- Add **user evaluation and suggestion** functions in the interactive interface.
- Hand over chat records with **low ratings**.
- **Analyze** user suggestions and **optimize** accordingly.

**User behavior tracking**
- Obtain the **dissatisfied dialogues** of users by analyzing the behavior data.
- Such as the **records before the user switches to the manual customer service**.
- The **chat records before user closes the conversation** for the problem is not resolved.

**Back-end model analysis**
Implementation journey
Most of models used in CAI requires strong computing power. Cloud computing is highly recommended to process massive amount of data simultaneously and then run algorithms to achieve the results, which can make up for the inefficiency of computer hardware.

The data of CAI includes configuration data and historical data. Both of them increase rapidly during the life cycle of product. Therefore, high-performance storage plan such as AI cloud should be considered at the beginning. Distributed-system is highly recommended for CAI.

Structured data used for training CAI includes the Q&A and free-style conversations, which should be collected and cleaned in advance according to business needs. Data governance rules should be established and the classification of data must be correct because all the data is highly confidential and is the foundation of CAI.

Unstructured data includes pictures, voices, and unorganized documents. All of them can be used in CAI. Related technologies such as Natural Language Processing, Speech Recognition, Optical Character Recognition and Machine Learning can be used to process these types of data.

CAI needs to be in accordance with the organization’s process of digital transformation, culture and strategy. Organizations’ structure should be adjusted according to AI strategy, including business empowerment and innovation, technology development, change management, etc.

This project requires the cooperation of employees from multiple departments and the project manager’s supervision. AI builders and AI translators are all of great importance. Companies will need the right mix of talents to translate business needs into solution requirements, build and deploy AI systems, integrate AI into process and interpret results.
CAI implementation journey

**Image**
- Initial vision
- Digital diagnosis
- Proof of concept
- Infrastructure

**Deliver**
- Evaluation of application & deliver a roadmap
- Iteration & update
- Customized optimization

**Run**
- Collect feedbacks from users
- Analyze feedbacks and logs
- Optimization
- Launch
- Deploy

**Goal**
- Deliver a roadmap
- Iteration & update
- Customized optimization

**Time**
- 2-4 Weeks
- 8-16 Weeks
- Continuous
## Talents required during the lifecycle of AI adoption

<table>
<thead>
<tr>
<th>Role</th>
<th>Responsibilities</th>
<th>Image</th>
<th>Deliver</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Program Leader</strong></td>
<td>Govern the initiative and project, manage key issues and risks, approve project scope and timeline changes.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Product Manager</strong></td>
<td>Schedule internal meetings, weekly status reports, manage risks/issue and escalations.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Architect</strong></td>
<td>Define standards of architectures and technology stack and lead the architecture design of the framework.</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td><strong>DBA</strong></td>
<td>Work with architect to design the data flow and manage the data used or generated by CAI.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Algorithm Specialist</strong></td>
<td>Develop the key algorithms or evaluate the algorithms from the vendors.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Software Developer</strong></td>
<td>Develop the CAI by integrating the algorithms to the CAI framework and optimize the product continuously.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Data Analyst</strong></td>
<td>Analyze the logs and users activities to support optimization and work with consumer behavior specialist to establish the analysis process.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Test Analyst</strong></td>
<td>Test the product including the logics, securities and data flow.</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td><strong>DevOps</strong></td>
<td>Build the test environment, help development team deploy the CAI and continuously monitor the operating status of the application.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Business Leaders</strong></td>
<td>Translate business problems/needs into requirements that guide the building of solutions, and to interpret results from CAI system and make decisions.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>UI Designer</strong></td>
<td>Design the UI of the product and make AI systems easier to navigate.</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td><strong>Marketing Specialist</strong></td>
<td>Design the schedule of the promotion and promote CAI in the organization or to the public.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Consumer Behavior Specialist</strong></td>
<td>Help the team establish the scope, design the interaction of the product and define the standard of the behavior analysis.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td><strong>Change Management Experts</strong></td>
<td>Implement change strategies and help integrate AI into the organization’s processes.</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Step</td>
<td>Step name</td>
<td>Details</td>
<td>Key consideration</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
</tr>
</tbody>
</table>
| 1    | Initial vision                    | • Evaluate the **situation** of the organization to identify the challenges and opportunities  
• Develop initial **high-level vision** of CAI, which is in line with the corporate value and strategy  
• Define the list of **application scenarios** to find the suitable solutions | • It is better to image the application of CAI from the **macro perspective**, such as corporate digital transformation strategy, corporate AI planning, etc. |
| 2    | Digital diagnosis                 | • Understand the current status of IT foundation  
• Understand the **data governance** status  
• Evaluate **talents** required for each procedures of this project | • Business leaders are required to **translate** business needs into solution requirements |
| 3    | Evaluation of application & Deliver a roadmap | • **Prioritize** application scenarios and evaluate IT foundation  
• Integrate information and make a related **roadmap** | • Key factors: IT Foundation, Data readiness, value delivered, project cost, time scope and etc. |
<table>
<thead>
<tr>
<th>Step</th>
<th>Step name</th>
<th>Details</th>
<th>Key consideration</th>
</tr>
</thead>
</table>
| 1    | POC       | • Establish the **baseline** of core models  
• Evaluate the **performance** of the algorithm and key functions | • POC is mainly to verify the **feasibility** of the algorithm |
| 2    | Infrastructure | • Test the computing power and capacity of the storage  
• Build a new database or integrated with existing data warehouse | • Applicability of the architecture  
• **Integrate** with the organization’s IT structure  
• **AI cloud** is recommended to operate and store data |
| 3    | Prototyping | • Establish and improve **interactive interface**  
• Confirm the correctness of data flow  
• Put core module into actual **business scenarios** | • Actual **business needs** should be evaluated  
• AI translators are needed to bridge technology and business |
| 4    | Test      | • Test **logic** of the algorithms and **accuracy** of the results  
• Application test such as **user friendliness** | • The algorithms need to be tested **repeatedly** to further optimize accuracy  
• Less **running time** can bring better experience to users |
| 5    | Iteration | • Update the models with the **feedback** from the step of test  
• **Optimize** the interaction according to users’ activities | • It is a **cyclic process**  
• **Continuous** testing leads to continuous optimization |
| 6    | Deployment| • Deploy CAI in production environment and connect it to production database | • **AI cloud** can be used to collect and store massive data of all users which is useful for optimization |
### Methodology | Run

**Objective**

Formally put CAI into use and optimize models continuously to make it better fulfill its original vision.

<table>
<thead>
<tr>
<th>Step</th>
<th>Step name</th>
<th>Details</th>
<th>Key consideration</th>
</tr>
</thead>
</table>
| 1    | Launch   | • Officially open for use  
      |          | • Training  
      |          | • Promotion  
      |          | • Feedback collection  | • Change management experts should improve employee’s acceptance of CAI, align employee’s goal with the organization’s vision  
|      |          |         | • Organization reconstruction and process optimization |
| 2    | Collect feedbacks from users | • Use event tracking to record the users’ activities  
      |          | • Collect the users’ reviews with the help of user feedback mechanism  | • Event tracking should be automated. Machine learning based classification is recommended to reduce labor cost  
|      |          |         | • Users’ activities and reviews should be combined to analysis comprehensively |
| 3    | Analyze feedbacks and logs   | • Determine the popularity of each function by analyzing the users’ activities  
      |          | • According to the users’ suggestions to find the directions for improvement  
      |          | • Evaluate metrics of core functions  | • The real needs can be better understood through analyzing user activity logs  
|      |          |         | • Additional workshops and questionnaires to address users’ requirements |
| 4    | Optimization | • Optimize the core models and related configurations to improve the accuracy of the module and make it better serve real life  | • Optimization is a continuous work at the life cycle of CAI and it is the key step to make a smarter CAI |
CAI configuration approach -- trade-off between precision and recall

The trade-off between precision and recall is an important topic of CAI. According to different requirements and scenarios, clients have specific restriction of CAI. Some scenarios require CAI to be as precise as possible, while others require CAI to maintain the dialogue as much as possible, constantly attracting users to engage.

Deloitte CAI five quadrants

**Recall**

- **Feature:** CAI responses users as much as possible but may be irrelevant.
- **Suit for:** Speak freely, can accompany users and have fun with them when queuing for example.

**Precision**

- **Feature:** High accuracy, and zero mistake tolerance.
- **Suit for:** Situation that needs carefully treated and of great importance, sometimes may be confidential to users, such as transactions.

- **Feature:** Relatively high accuracy and can give highly relevant answers or actions.
- **Suit for:** High relevance requirement and high response effectiveness.

- **Feature:** Highly relevant answers can be achieved in most cases.
- **Suit for:** Some task-oriented work and interactions in life such as housekeeper robot.

- **Feature:** Use “I don’t know” to replace some uncertain answers.
- **Suit for:** Assistant for problem solving which can do a pre-talk before service officers’ help.

**CAI configuration approach**

- **Companion AI**
- **Warm up AI**
- **Customer service**
- **Guide robots**
- **Smart home**
- **Phone assistant**
- **Game AI**
- **AI outbound**
- **Shopping guide**
- **Smart car**
- **RPA chatbots**
- **Educational CAI**

**Deloitte CAI five quadrants**

- **Recall**
  - Companion AI
  - Warm up AI
  - Customer service
  - Guide robots
  - Smart home
  - Phone assistant

- **Precision**
  - Game AI
  - AI outbound
  - Shopping guide
  - Smart car
  - RPA chatbots
  - Educational CAI
Technology of precision and recall

In order to achieve the five quadrants, related mechanisms are recommended to make it more suitable for use, which are broken down into three aspects: algorithm models, rules and training data.

**Model**

- **Generation Based**
  - New replies can be synthesized from massive data
  - Answers based on the language ability learned from huge amount of conversations

- **Retrieval Based**
  - Responses are highly relevant
  - Can only return the answer which exists in the knowledge base

**Rules**

- **Lower**
  - Loosen the restriction rules of answers to enrich the diversification of replies

- **Threshold of relevance**
  - Standardize the format of users’ questions to match precise answers in knowledge base

- **Higher**

**Data**

- **Massive**
  - Real conversation data
  - Encyclopedia
  - Newsletter & magazines

- **Specific**
  - Regulations & rules
  - Business & finance terminology
  - Medical jargons

Recall

Precision
D.Bot introduction
D.Bot | Intro

Why we design D.Bot?

- After providing professional services and solutions to clients for many years, Deloitte has found how to minimize repetitive human tasks, improve user experience and prepare for the future of work, which are some of the pain points clients are facing today. Many clients wish to have interactive methods to realize their digital transformation and keep their competitive edges.
- D.Bot is born from these pain points and incorporates common requirements from different clients. At present, D.Bot has become a representative product of CAI designed by Deloitte.

What can D.Bot bring?

Efficiency improvement
- Reduce communication cost
- Reduce labor effort
- Reduce language complexity
- Increase productivity
- Increase operational efficiency
- Increase interactive friendliness

User-centric
- D.Bot aims to provide a user-friendly engine, helping clients deploy customized chatbots quickly

Promote future work
- As a CAI engine to assist employees efficiently
- As an AI robot to respond timely
- As an interpreter to change the method of interaction

What is behind D.Bot?

A combined CAI technology

Key functions
- Machine learning
- NLP
- Voice processing
- FAQ
- Task-oriented
- Machine comprehension

D.Bot | Intro
# D.Bot | Function

D.Bot incorporates the most common functions of CAI. All the models of these functions are designed with new architecture and algorithms and trained with massive dialogue data.

<table>
<thead>
<tr>
<th>Description</th>
<th>Technology</th>
<th>Scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FAQ</strong></td>
<td><strong>Text semantic matching</strong>: SimNet is used in FAQ module. It is a point-wise model which applies the method of representation-based match.</td>
<td><strong>Q&amp;A Robots</strong></td>
</tr>
<tr>
<td><strong>Task-oriented</strong></td>
<td><strong>Conversational deep learning</strong>: This module uses deep neural network and rules to complete tasks.</td>
<td><strong>Task management</strong></td>
</tr>
<tr>
<td><strong>Machine comprehension</strong></td>
<td><strong>Attention mechanism</strong>: It is a joint model based on attention mechanism, using different outputs of the model to complete task classification and slot labeling prediction in the same model.</td>
<td><strong>Q&amp;A related to laws and legal requirements</strong></td>
</tr>
</tbody>
</table>

### FAQ
- **What it is**: FAQ module builds dialogue capabilities through Q&A pair. It mainly supports a single round of knowledge-based dialogue.
- **How it works**: FAQ module needs users to prepare question and answer pairs in advance and it will generate answers from these pairs.
- **Pros and cons**: High accuracy, Low interactivity.

### Task-oriented
- **What it is**: The task-oriented module is a multi-round dialogue system.
- **How it works**: It supports the task which is users come with a clear purpose hoping to get information or services that meet certain restrictions. As the users’ purpose can be complicated, it may need to be presented in multiple rounds and users may modify or perfect his purpose during the dialogue. In addition, it also requires the module help users get good response by asking, clarifying or confirming.
- **Pros and cons**: High accuracy, high interactivity.

### Machine comprehension
- **What it is**: The machine comprehension module is an search robot which can find answers in large documents according to the multiple questions.
- **How it works**: Machine comprehension module is applicable to the scenario where the client has multiple knowledge documents and the answer can be generated by directly intercepting the content of the documents.
- **Pros and cons**: The advantage of the module is that it does not require clients to maintain Q&A pair which solves the last mile problem. The disadvantage is that the accuracy is lower than FAQ module and the response speed is relatively slower.

- **Dual module**: The module has two core models which are paragraph extraction model and machine comprehension model.
- **Coordinated mechanism**: The first model helps the module extracts the candidate paragraphs which may have potential answers. The second model extracts answer fragments from the candidates.
- **Massive training**: Both models are pre-trained with tens of thousands of documents.
D.Bot | Cognitive technology

D.Bot involves combination of traditional machine-learning algorithms and deep neural network.

- Traditional machine-learning algorithms are used to accelerate inferencing and get interpretable results.
- Deep neural network is used to build a generalized model and obtain more accurate results.

Cognitive solutions: main capabilities

- Voice recognition
- Speech synthesis
- Multilingual processing
- Text classification
- Text clustering
- Dialogue mgmt.
- Machine learning
- Natural language processing
- Slot filling
- Text similarity
- Entity recognition
- Machine comprehension
- Voice processing

Cognitive enabler

- Natural Language Processing
  - Word to vector
  - Support vector machine
  - Dialogue state tracking
  - Finite state machine
  - Linear discriminant analysis
  - Inverse document frequency
- Voice Processing
  - Recurrent neural network
  - Latent semantic indexing
  - BERT
  - Memory network
  - DSSM-LSTM
  - Conditional random field
- Machine Learning
  - Voice Activity Detection
  - Sequence to Sequence
  - WaveNet / Tacotron
  - Hidden Markov Model
D.Bot | Architecture

### Interaction layer
- **Connected to daily-used IM software easily**
-  D.Bot provides services through RESTful API
-  Developer can easily connect it to **various IM platform**, such as WeChat official account, Skype and web assistant
-  It greatly reduces the difficulty of **development** and increases the efficiency

### Function layer
- **Different functions meet different requirements**
-  D.Bot uses modular design to allow clients to **select functions according to requirements**
-  All functions are **user-friendly** which can be configured and **customized** quickly, even if the vision changes
-  Clients can remove or add a function without affecting the operation of other functions

### Analytics Layer
- **Real-time supervision and customized development for analytics**
-  D.Bot provides a strong module of analytics to **optimize CAI** during the product life cycle
-  It contains a dashboard, which monitors the operating status and presents the statistical result of historical data
-  It also has a toolkit, which has multiple **machine learning** and **statistics tools** to help data scientists and business specialists to analyze users’ activities

### Storage layer
- **Mass storage and excellent database integration**
-  Storage layer constructs the foundation of D.Bot, which stores the knowledge of CAI and **continuously** records operating status
-  It greatly reduces the cost of migration and maximizes the **utilization rate** of existing data in the organization

### Interaction layer options
- WeChat
- Skype
- Webinar
- Web
- Email

### Function layer options
- FAQ
- Task-oriented
- Machine comprehension
**D.Bot | Industries**

The vision of D.Bot is to create the CAI engine that will be applied across industries. Therefore, when Deloitte designed D.Bot, the demands in various industries were taken into consideration.

**Scenario:** D.Bot can be used as a customer service assistant online. D.Bot can recommend products to consumers by capturing user’s requirements during the dialogue.

**Benefit:** D.Bot is a good tool to increase consumer satisfaction with customer service and reduce the labor cost in post-sales and pre-sales period.

**Scenario:** D.Bot is a good assistant in logistics tracking. It is a supplementary toolset to improve internal management.

**Benefit:** CAI application plays an important role in digital transformation, helping clients manage their digital assets efficiently.

**Scenario:** Smart customer services and smart sales are two main scenarios in financial services.

**Benefit:** The main benefit of the application is greatly reducing the labor cost of repetitive work.

**Scenario:** D.Bot enhances e-government and smart cities. It has advantages in communicating with citizens.

**Benefit:** D.Bot’s analytics module can help governments understand requirements of residents more clearly and quantitatively evaluate the level of governance.

**Scenario:** Smart healthcare is a developing direction of life science. D.Bot can be used in the process of diagnosis, guidance, management of digital medical records.

**Benefit:** It improves the satisfaction of patients and liberates doctors and nurses to focus on high-level work, which alleviates the shortage of medical resources to a certain extent.

**Scenario:** D.Bot can be connected to various IM software, it can be used on different media as an assistant.

**Benefit:** D.Bot can reduce the entry cost of CAI, including labor costs and time costs.
Deloitte’s CAI vision and aspiration

Design the future of work

Reshape the future of life

Break the barrier of language

Connect human and machine
Joseph Chu is the Chief Digital Officer (CDO) of Deloitte China. He has more than 25 years of experience in digital transformation, technology advisory, big data and artificial intelligence. As Deloitte AI Institute leader in China, he is focusing on leveraging artificial intelligence technologies such as facial recognition, natural language processing, robotics, etc. to drive business value. He is leading a digitization task force that is comprised of senior business leaders and AI experts to develop our digital and cognitive strategy and deliver our AI enabled services to the clients.

Deloitte AI & Cognitive Service Team

Chief Digital Officer
Deloitte China

Deloitte AI Institute (DAI) is Deloitte China’s digital technology research team. It has nearly 30 AI experts and is responsible for applying AI and cognitive technology to Deloitte client solutions and services.

Innovation & Digital Development Center
Innovation & Digital Development Center (IDDC) is responsible for building Deloitte’s digital assets and most of the assets will be enabled by AI and cognitive technologies.

Audit & Assurance
- Audit Innovation Team
- Assurance Digital Team
- Robotics & Cognitive Automation Team

Consulting
- Analytics & Information Management Team
- Digital Team

Risk Advisory
- RA Assurance Team
- RA Analytics Team

Financial Advisory
- Analytics & Digital Innovation Team
- FA Innovation Team

Tax & Legal
- Robotics & Cognitive Automation Team

30+
AI enabled Assets

2300+
AI professionals

30%
Revenue enabled by AI / Cognitive

© 2019. For information, contact Deloitte China.
Deloitte is a leading global provider of audit and assurance, consulting, financial advisory, risk advisory, tax and related services. Our global network of member firms and related entities in more than 150 countries and territories (collectively, the "Deloitte organization") serves four out of five Fortune Global 500® companies. Learn how Deloitte's approximately 312,000 people make an impact that matters at www.deloitte.com.

Deloitte Asia Pacific Limited is a company limited by guarantee and a member firm of DTTL. Members of Deloitte Asia Pacific Limited and their related entities, each of which are separate and independent legal entities, provide services from more than 100 cities across the region, including Auckland, Bangkok, Beijing, Hanoi, Hong Kong, Jakarta, Kuala Lumpur, Manila, Melbourne, Osaka, Seoul, Shanghai, Singapore, Sydney, Taipei and Tokyo.

The Deloitte brand entered the China market in 1917 with the opening of an office in Shanghai. Today, Deloitte China delivers a comprehensive range of audit & assurance, consulting, financial advisory, risk advisory and tax services to local, multinational and growth enterprise clients in China. Deloitte China has also made—and continues to make—substantial contributions to the development of China's accounting standards, taxation system and professional expertise. Deloitte China is a locally incorporated professional services organization, owned by its partners in China. To learn more about how Deloitte makes an Impact that Matters in China, please connect with our social media platforms at www2.deloitte.com/\cn/\en/social-media.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited ("DTTL"), its global network of member firms or their related entities (collectively, the "Deloitte organization") is, by means of this communication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser.

No representations, warranties or undertakings (express or implied) are given as to the accuracy or completeness of the information in this communication, and none of DTTL, its member firms, related entities, employees or agents shall be liable or responsible for any loss or damage whatsoever arising directly or indirectly in connection with any person relying on this communication. DTTL and each of its member firms, and their related entities, are legally separate and independent entities.

© 2020. For information, contact Deloitte China.