

Company: IBM, Armonk, NY, USA
Company Description: IBM is the hybrid cloud and AI technology and services leading provider, focused on delivering client value through a combination of technology and business expertise. IBM solutions draw from an industry-leading portfolio of capabilities in software, consulting services and a deep incumbency in mission-critical systems, all bolstered by one of the world’s leading research organizations.
Nomination Category: Technology Categories
Nomination Sub Category: Best Technical Support Solution - Computer Services
Nomination Title: Build Once, Use Anywhere: - Scaling Enterprise AI



1. Which will you submit for this nomination, a video of up to five (5) minutes in length or a written essay of up to 650 words? Choose one:

Essay of up to 650 words
2. If you are submitting a video of up to five (5) minutes in length, provide the URL of the nominated video here, OR attach it to your entry via the "Add Attachments, Videos, or Links to This Entry" link above, through which you may also upload a copy of your video:
3. If you are providing a written essay for this nomination, submit in this space an essay of up to 650 words describing the nominated Technical Support Solution since 1 January 2023:

Total 650 words used.

The 2024 has been an amazing year of transformation for IT departments, with all teams realizing the power of AI and Large Language Models (LLMs). The IBM Cognitive Tooling Team within the IBM CIO is a small, 28 member team that foresaw the power of LLMs and made it their strategy to not only build its own AI solutions with LLMs, but also build with frameworks that are used by many other IBM teams to provide their own AI solutions for their user base. This "build once, use anywhere" vision enables teams without any much experience in AI to be able to deploy their own solutions with ease.

IBM Cognitive Tooling created two frameworks for extended IBM teams to use to quickly spread AI solutions across the enterprise at scale. Each include metric collection processes to enable all onboarded teams to be able to measure the impact of their AI solutions.

1. The Unified Chat Framework (UCx) is the IT vision to provide a chat experience that can be implemented almost anywhere on our webpages as well as business applications (ie, Salesforce). UCx provides the entire technical setup and tools that enable the chat experience with all functionalities. This means all teams that onboard onto UCx can focus on what they do best: using their expertise to provide the experience and content they need to serve their user base. The framework includes each of the following configurable features:
 1. Business unit specific conversation flows
 2. User Personalization
 3. Automation of tasks to enable the user self-serve
 4. Intelligent Routing: With nearly 100 onboarded teams so far this year, the chat solution uses AI to intelligently determine the best solution to present to our users.
 5. Generative AI: Using a Retrieval Augmented Generation (RAG) approach, UCx enables each IBM onboarded team to provide a fine-tuned response for each user query. In addition, a hallucination detection system for GenAI responses using Quantum computing ensures the generated answers are grounded in the content, so they are indeed correct.
 6. The ability to route from one experience to another: For example, if a user comes into the marketing chat experience, but has a product support question, the chat can automatically take the user to the correct support experience to best serve what the customer needs.
 7. Live chat with a human: If a user requires contact with a human, the Unified Chat framework can route the user to the correct support agent, sales rep, etc.

Overall, UCx has been onboarded by 96 different teams in IBM across sales, marketing, customer support, training, and business partner organizations. Over 65,000 chats have been initiated through Unified Chat (YTD 2024). The chat volume has grown 429% in the last year. We have measured an average of 74% containment for the chat experience, which means the users have been able to self-serve their own needs through the chat experience. The experience currently exists on over 60,000 pages across IBM.

1. Magma is a "build once, use anywhere" framework built by the Cognitive Tooling team to accelerate the creation of AI applications by providing trusted components that simplify AI implementations for all IBM teams. Teams are empowered to focus on the novelty of their ideas rather than on addressing technical dependencies of implementing AI solutions.

The Framework provides customizable toolchains which are sequences of steps that can perform tasks (Q&A, Summarizations, etc) using generative AI. It also provides prebuilt connections to a myriad of IBM data sources.

While still in its early stages MAGMA is used by 16 different teams/use cases across IBM to process over 2.1 million AI transactions YTD. As an example of only one of these use cases, MAGMA is used to produce customer support case summaries, which have saved 630K hours by IBM support agents. It is also used in the generative AI for the Unified Chat solution.

Overall, the team enabled \$214M in business benefits by increased productivity and automation.

4. In bullet-list form (up to 150 words), provide a brief summary of up to ten (10) of the chief achievements of the nominated Technical Support Solution since 1 January 2023:

Total 150 words used.

- The IBM Cognitive Tooling Team’s foresight into building reusable AI Frameworks for IBM teams to more easily deploy AI solutions has been used by 112 different IBM teams deploying personalize AI solutions for their user bases.
- The team is a small group of 28 Developers, Data Scientists, Architects, Testers and Business Analysts works as an incubation hub to help other IBM team do what they do more effectively.
- The overall benefit on the team’s deployed solutions for 2024 YTD across the key business units of Customer Support, Sales, and Marketing are a combined \$214M. The solutions combine to in business benefit by increased productivity, assists in better decision making, and deflection / avoidance of work through automation, insights, and AI.

Included in the Supporting materials are some of the external facing chat experiences created using the UCx framework. Just click on the chat icon on the bottom right of the page.

Attachments/Videos/Links:

[Build Once, Use Anywhere: - Scaling Enterprise AI](#)

[REDACTED FOR PUBLICATION]