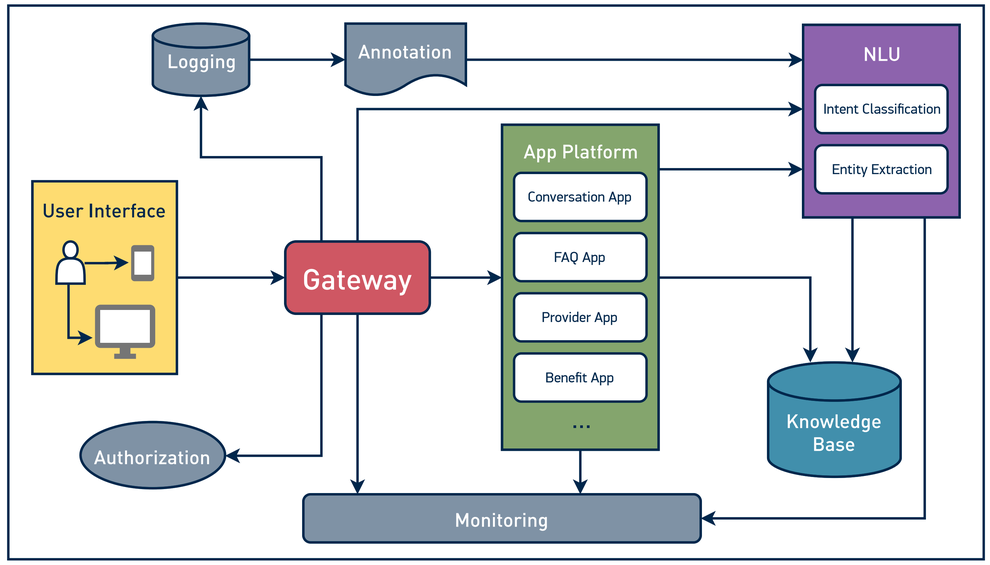
A Multi-Task Healthcare AI Bot Platform

Martin Horn, Xiang Li, Lin Chen, Sabin Kafle

Cambia Health Solutions
Seattle, WA 98101

Correspondence: Martin.Horn@cambiahealth.com



Summary

Healthcare is a complicated and confusing system, often placing a large workload on customer support services.

Conversational agents present a very useful application of facilitating and optimizing the interactions between the customers and support staff, with bots answering simple questions while deferring complex individualized questions to the support staff in a timely fashion.

AI bots are developed either through a task-oriented approach based on the domain requirements or an end-to-end data driven approach, which is primarily used to develop social bots.

However, there are a certain set of challenges for such a product in the healthcare domain, including confidentiality and domain knowledge. In healthcare it is more useful to have a modular task-centric design for both flexibility and compliance.

We have built a secure AI bot platform for the healthcare domain with the hope to overcome some of these obstacles and drive progress on practical, useful conversational agents.

It has been designed to be able to learn more intents and functionalities over time, connect to more services, and wear different personas for different applications and audiences.

TOOLS & RESOURCES

Highlights

There are special constraints on industrial healthcare software systems, such as HIPAA compliance and domain knowledge.

Building an AI bot requires the creation of many similar components and skills.

A modular microservices architecture allows components to be re-used and independently secure.

Templates can enable easier bot skill development and NLP model training on small labeled datasets.

Performance & Evaluation

Table 1. Intent Classification Scores

Intent	P	R	F1	#
auth	0.93	1.00	0.96	25
copay	0.89	0.96	0.92	25
schedule	0.89	0.96	0.92	25
coinsure	0.85	0.88	0.86	25
...	...	...	...	...
cost	0.92	0.62	0.74	71
glossary	0.74	0.71	0.72	41
faq	0.44	0.83	0.57	35
general	0.65	0.50	0.57	26
Avg / Total	0.83	0.81	0.81	747
Overall Accuracy				0.81

Table 2. Entity Extraction Scores

D	Entity Type	P	R	F1
P	city	0.96	0.93	0.95
	member_id	1.00	1.00	1.00
	network_status	1.00	0.97	0.99
	facility	0.80	1.00	0.89
	practitioner	0.95	0.92	0.93
	specialty	0.96	0.90	0.92
	state	0.98	0.98	0.98
	zip_code	1.00	1.00	1.00
B	benefit_category	0.95	0.85	0.90
	member_id	1.00	0.98	0.99
G	concept	0.94	0.89	0.92

Traditional classification metrics on Natural Language Understanding models:

- **Table 1**: Precision, recall, F1 score, micro-averages, and overall accuracy for MaxEnt intent classifiers on bottom-tier intents
- **Table 2**: Precision, recall, F1 score for CRF entity extractors for each entity type using strict entity-level matching, grouped by (D)omains: (P)rovider, (B)enefits, (G)lossary

During runtime, we monitor service metrics like job health, response latency, and queries per second.

We collect feedback and calculate custom usage statistics via our logging service.

Bot responses for skills like Provider Search are validated by in-domain experts.

Modular Design

- Standalone Docker-containerized gRPC services using JWT-based authentication
- Services can be re-used for different purposes or by different clients
- Strict authorization ensures each service is secure and compliant in a sensitive environment containing Protected Health Information (PHI)
- We use the concept of a “persona” to allow configurations that turn on different skills or personalities for a specific product or brand

The Pipeline

User Interface: desktop UI for in-house bots and mobile UI for consumer facing healthcare app

Gateway: connects to each microservice—including authorization, monitoring, and logging services—and directs traffic

App Platform: extensible platform which houses each app/skill with its own dialog management and business logic

Knowledge Base: knowledge store for information around providers, members, medical terminology, benefits, pharmaceutical drugs, etc.

Natural Language Understanding: hierarchical classifier for determining user intent and sequence labeling classifier for extracting entities from query

Conclusion

- We built an AI bot architecture consisting of upgradable modules which form the foundation for building a stable AI bot with value to customers.
- Our system consists of a flexible language understanding module, an app platform facilitating dialogs and performing tasks, robust knowledge base services, and a gateway connecting each microservice together.
- The modularity of the system aids its flexibility, extensibility, and compliance in the sensitive healthcare domain.
- In future work we would like to integrate a more expansive knowledge graph to further empower the AI bot in handling complex, personalized, domain-specific queries.