Cognitive Services and Intelligent Chatbots: Current Perspectives and Special Issue Introduction

Amit Sheth
Wright State University

Hong Yung Yip
Wright State University

Arun Iyengar
IBM TJ Watson

Paul Tepper
Nuance Communications

■ **DURING THIS DECADE,** Artificial Intelligence (AI) has established itself as a phenomenon in our daily life. It is one of the most important investment priorities (http://bit.ly/Invest-AI) and has been at the forefront of recent technological disruptions. New boundaries are being pushed daily, ranging from winning against the world's best Go player (http://bit.ly/Alpha-Go) to self-driving and autopilot capabilities (http://bit.ly/Autopilot-Drive). Startups and tech giants continue to forge ahead in applying AI in diverse domains spanning speech recognition (http://bit.ly/Speech-AI), image classification (http://bit.ly/Visual-Recognition), recommendation systems (http://bit.ly/Recom-Engines), and internet search (http://bit.ly/Internet-Search). Colloquially, AI can be loosely described as a machine mimicking cognitive abilities that humans exhibit fundamentally such as learning and problem solving.[1] Numerous breakthroughs such as the Google Duplex (http://bit.ly/GDuplex), an AI assistant that is capable of impersonating verisimilitudinous phone conversations, have led to praise as well as incredulity among technology enthusiasts. Nonetheless, advances in computing, communication, social-interaction, and Web technologies, as well as embedded, fixed, and mobile sensors and devices with internet presence - commonly referred to as Internet of Things (IoT), are redefining the way that humans interact with computers.

MEDIA RICHNESS THEORY

Media richness theory introduced by Daft and Lengel in 1986 is a framework used to describe, rank, and evaluate a communication medium's ability to reproduce the information richness sent over it.[2] Put simply, it is the ability of a medium to handle multiple information cues simultaneously and facilitate rapid feedback to establish a personal focus.[3] With the current rate of consumers' acceptance of wearable technology[4] and the proliferation of low cost IoT devices (Figure 1), it is now possible to capture a multitude of media-rich information and integrate them to deliver far more effective communication while augmenting reality.

Such unprecedented growth of IoT, interaction devices, and mobile computing adoption calls for intelligent data-driven computation and decision-making frameworks to abstract these heterogeneous yet complementary raw sensor data into actionable information and meaningful knowledge. Recent years have seen tremendous growth

01	Webcam/ Wearable Cameras	Visual social cues of gestures, facial and body language as well as emotions (http://bit.ly/Azure-Emotion)
02	Microphone/ Headsets	Vocal cues of emotions by physical properties of voice such as vocal intonations (http://bit.ly/Empath-Vocal)
03	Smartwatch/ Activity Tracker	Heart rate, activity monitor, sleep hygiene, Global Positioning Satellite (GPS) motion, altitude, vibration, and peripheral capillary oxygen saturation (SpO2) data (http://bit.ly/Fitbit-Track)
04	Smart Glasses	Audio and visual, Augmented Reality (AR) (https://www.x.company/glass/)
05	Smart Clothing	Electrocardiography (ECG), heart rate variability, breathing rate, minute ventilation, activity intensity, peak acceleration, steps, and sleep positions (https://www.hexoskin.com/)
06	Air Quality Monitor	Air quality sensing such as fine particles, total volatile organic compounds, temperature, and humidity (https://foobot.io)
07	Motion Detector	Movement

Figure 1. Sensor and wearable technology with their respective modality of information.

in AI research which have in part, seeded, and propelled the emergence of various cognitive services as mediatorial dominion between AI and IoT data to harness and unfold the information richness to achieve media naturalness and synchronicity.

A chatbot (also known as a conversational agent or virtual assistant) on the other spectrum, has also became increasingly popular due to its capability of simulating human-like conversations with a user through speech, text and multimodal communication (http://bit.ly/Chatbot-Media). In fact, an interesting trend is that chatbot-assisted queries are 200 times more conversational than search, and users are demanding more human-style interaction (http://bit.ly/Chatbot-Trend). However, most of today's chatbots do not truly understand natural language, nor do they have cognitive capabilities to understand the context of their conversations, like world knowledge or commonsense reasoning. This means that they lack the ability to go beyond scripted conversations to make interactions with computers feel like natural conversations. Rapid progress in conversational AI facilitated by continuing advances in machine learning (ML) and natural language processing (NLP), and other cognitive services may usher in the next generation of these systems, enabling them to move beyond simple, scripted

conversations and towards richer conversational interaction. What started as speech focused chatbots are rapidly being integrated with smart displays that are sure to add visual component to the next generation of chatbots. The next section describes how various online cognitive services can be integrated with chatbots to substantially enrich human–computer conversational experience, exemplified with a few examples of use cases in the healthcare domain.

BRIDGING CHATBOTS WITH ONLINE COGNITIVE SERVICES

Online cognitive computing or cognitive services, broadly speaking, are services (usually web or cloud hosted) that incorporate AI including but not limited to knowledge graphs, NLP, and ML techniques. Cognitive computing can be loosely described as the ability to simulate human thought process in a computerized model. It is a higher level of functional API that takes in raw unstructured data in the form of text, speech, images, videos, and in future other modalities, and converts them into smart data[5] through means of complex computations often including AI techniques trained on big data. With the availability of abundant compute and storage resources owing to cloud computing,

combined with the evolution of analytics, cognitive computing has seen tremendous growth in crowd adoption as it becomes more affordable and accessible to consumers and businesses. IBM Watson Services and Microsoft Cognitive Services are two representative cognitive computing platforms that provision powerful capabilities through simple web APIs. Although they were initially designed for developers, both companies are continuously building extra layers to ease into and align with enterprise needs.

Cognitive Services can be broadly grouped into five categories: language, speech, vision, knowledge, and search (http://bit.ly/Cog-Services). Examples from each category include: 1) language services are named entity recognition and linking, sentiment analysis, and intent classification; 2) speech services such as speech-to-text and vice versa; 3) vision services for example face recognition and automatic image captioning; 4) knowledge such as data insights and news discovery; and 5) search such as autosuggest, image, news, video, and web search.[6,7] With a collection of over 30 cognitive services APIs (http://bit.ly/Chatbot-APIs), the challenge, however, is to semantically piece together and address how an application, specifically a media-rich chatbot, can be made more intelligent when given access to a continuous influx of voluminous, streaming, and dynamic real-time sensor and multimodal IoT data. By intelligent, we mean the ability of a chatbot to contextualize (interpret data in a personal context with respect to domain-specific knowledge), personalize, and abstract data into information (asking the right questions at the right time with supporting decision making and actions informed from data),[8] and incorporate more human-like behavior such as feelings, emotions, and empathy. Figure 2 shows how a basic chatbot can be extended with various online cognitive services to achieve human-like intelligence. While IBM's Watson Cognitive services are shown in these examples, growing suites of alternative services are now available.

Major tech companies (Amazon, Facebook, and IBM) have embarked upon the use of voice technology. While success so far is modest, engaging voice conversations (rather than text) in chatbot hold a significant promise for future man-machine interaction. While the contemporary general-purpose chatbots have had limited success primarily due to brittleness in modeling context and ability to use broader knowledge, chatbots in a domain such as the *healthcare domain* hold significant promise and are receiving correspondingly high interest. This is in part due to ability to use multimodal data and access to structured knowledge of medicine, as exemplified by the digital personalized health applications for asthma and cardiovascular diseases.[9,10] With the use of IoTs extended with a suite of cognitive services as illustrated in Figure 2 to provide the surrounding context, chatbot conversations can be made more robust and meaningful. The following section illustrates how online cognitive services can leverage on multimodal data to provide access and integration with healthcare domains.

ACCESS AND INTEGRATION WITH HEALTHCARE DOMAINS

While feeding raw data generated from IoTs into various online cognitive services pipelines to translate them into meaningful insights is a step closer to intelligent chatbots, we believe that it is also important for the system to combine, demystify, and make sense of this data through contextualization, personalization, and abstraction. *Contextualization* refers to data interpretation in terms of knowledge (context). It usually refers to mapping fine-grain data covering various facets by determining the data type and value, and then situates it in relation to other domain concepts, thereby deriving a meaningful interpretation of results. As an example, a contextualized chatbot with domain knowledge can understand slang terms that are commonly used in social media such as "bupe" which refers to its medical jargon "buprenorphine." *Personalization* refers to the future course of action by taking into account various factors with respect to a particular patient such as health history, physical characteristics, environment, activity, and lifestyle. For example, a contextualized and personalized chatbot can tell the weather with regards to an asthmatic patient vulnerability to high ragweed pollen level and suggest appropriate actions. *Abstraction* is a computational
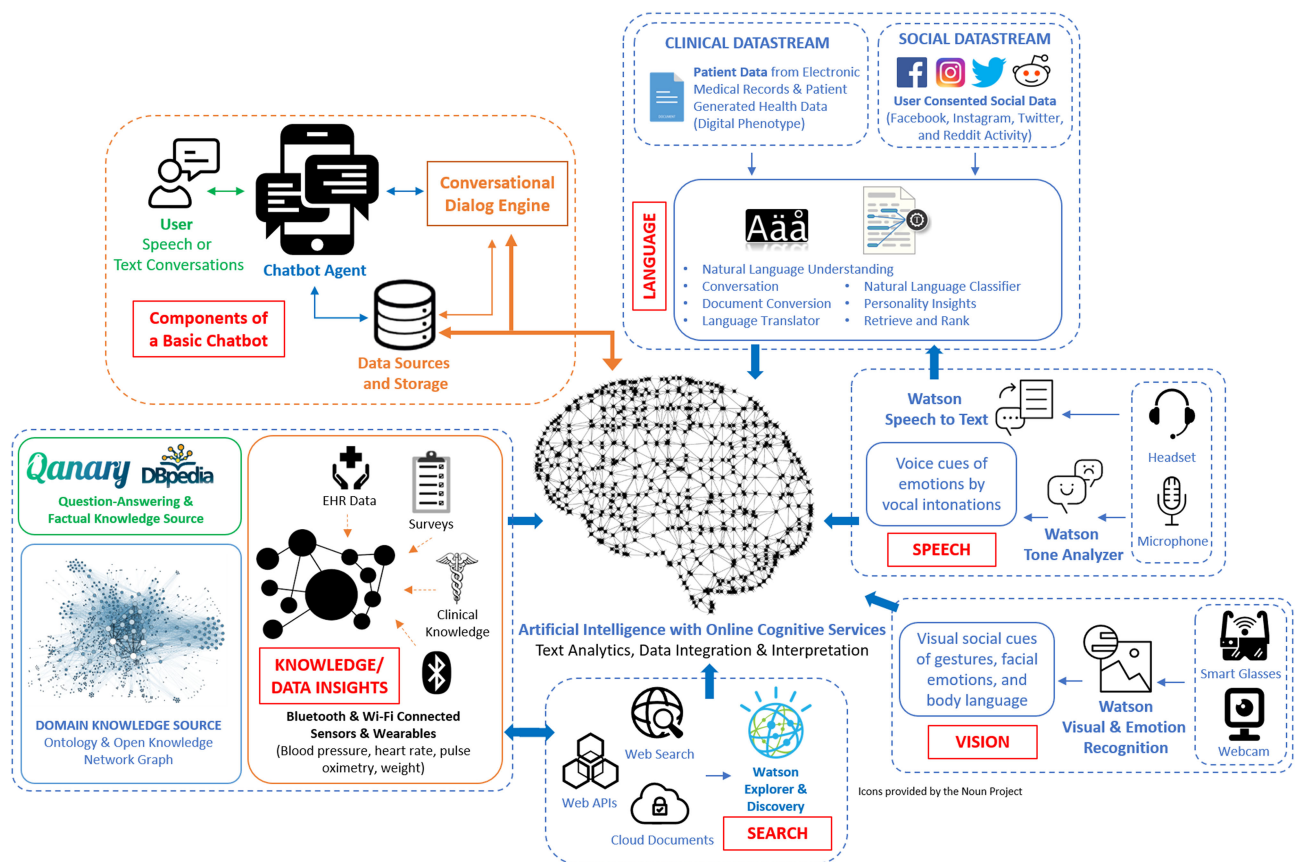
Figure 2. Illustration of how a basic chatbot can be extended using an example suite of cognitive services.

technique that maps and associates raw data to action-related information to provide an integrated view of proper remediation measures. For instance, high daily activity in the context of health can be abstracted and translated to a low risk of heart problems depending on demographic and genetic information as well as diet. To paint a more granular perspective, the following illustrates a few use-cases as to how such intelligent chatbots that leverage online cognitive services can be used as a "Personal Health Coach" in the healthcare domains.

As instances of intelligent chatbots built using cognitive services, we describe three intelligent chatbots under development in NIH funded projects on depressing/mental health (http://bit.ly/depression-social), pediatric asthma (http://bit.ly/kHealth-Asthma), and eldercare. In the case of mental illness, conventional screening often administers Patient Health Questionnaires (PHQ-9) to assess a patient's

depression severity. However, such screening has two inherent flaws. It relies heavily on a patient's ability to recall events that occurred over the span of the last two weeks, and all PHQ-9 criteria are given equal weight. For example, the criteria "feeling tired or having little energy" is weighted similarly to the "thoughts that you would be better off dead, or hurting yourself in some way." An intelligent chatbot has the ability to leverage various IoT devices and online cognitive services such as a camera to assess a patient's current mood via visual recognition; the microphone leveraging tone analyzer service to analyze the speech tonality and sentiment; personality insights service to determine the persona of the patient; access to patient's consented profiles to discover behavioral manifestations and slurs on various social media platforms; language understanding and translator services for linguistic analysis such as the use of slang terms; and QnA maker to formulate intelligent questions. A combination of these

online cognitive services can overcome the transitory nature of memory recall and capture subtle behavioral changes that are otherwise missed or not evidently translated during PHQ-9 assessment. These further allow a viable entry for psychotherapy delivery in accordance to the cognitive-behavioral therapy[11] and initiate the need for treatment intervention conforming to medical protocols.

In the case of asthma disease, traditional diagnosis such as the asthma control test is generally focused and constrained on information available from the patient to the doctor at the time of hospital visits. There are only so many insights that can be gleaned from such limited data. An intelligent chatbot, on the contrary, is capable of filling and bridging this information gap. For example, it can leverage online services for weather data and electronic medical records (parsed using cognitive services such as document conversion, natural language classifier and understanding) to understand patient susceptibility to ragweed pollen, combine these individual types of information, and suggest appropriate courses of actions. Imagine two different responses for the same query seeking weather information, "You can expect fairly sunny weather today" and "You can expect fairly sunny weather today, however the ragweed pollen level is a little high which does not look good for your asthma condition. Do minimize any outdoor activities." The former is an information retrieval task without cognitive ability, whereas the latter represents contextualized reasoning, inference, and recommendation capabilities which can enrich the patient's quality of life.

In the case of elderly care, aside from residing at the higher risk of developing chronic diseases, elder residents often fall into the technological skeptic and illiterate groups as compared to their younger compatriots (http://bit.ly/Elder-Technology). A chatbot provides a lower barrier of entry to elders who embrace the use of technology. As an orthogonal example, a chatbot equipped with speech recognition, multilingual translation, and natural language understanding cognitive services allows elders with literacy issues to either text or talk, whichever they feel more comfortable with. Alternatively, a chatbot with access to geographical,

locality, and Web APIs is also capable of providing social support such as organizing telehealth sessions, scheduling doctor appointments, and coordinating as well as arranging transportation services for elderly people with physical disabilities and transportation barriers, especially in congested cities. Nonetheless, such illustrated capabilities and functionalities have only scratched the surface of the broader possible use cases of online cognitive services for intelligent chatbots. Its emerging impact and seamless integration in healthcare delivery for augmented personalized health[12] such as self-monitoring, self-appraisal, self-management, intervention, and projection of disease risk and progression should be advocated and leveraged, befitting the best interests of both patient and physician.

IN THIS ISSUE

We have selected four papers that met high-quality standards for this special issue. These papers described the various dimensions of an intelligent chatbot, ranging from dialogue management to real-world applications and how it can bring harm if inappropriately abused.

In *"Approaches for Dialogue Management in Conversational Agents,"* Jan-Gerrit Harms *et al.* surveyed the field of dialog management and established an overview of the ways in which dialog has been approached. The authors then taxonomized, compared, and contrasted various dialog management tools including hand-crafted (rule-based), probabilistic (statistical), and hybrid approaches on a set of predefined dimensions of analysis such as dialog structure, runtime learning, error handling, dependencies, control, domain independence, and tool availability. The authors concluded that, despite the current state-of-the-art approaches, there are still ways to improve these existing dialog manager tools to realize the potentials as seen in their fictional counterparts such as data integration, context awareness, and policy generation.

Photo sharing has become an essential medium of communication and contemporary conversational assistants have made the communication between users richer and more convenient by enabling users to search and share their photos. However, such feature bears a tendency of privacy leaks and possibly high latency

between server (bot intelligence for photo mining and analysis) and smartphones communication (in-device photos) due to the conventional client-server architecture. In *"meChat: In-device Personal Assistant for Conversational Photo Sharing,"* Kang-Min Kim *et al.* proposed *meChat,* a novel in-device personal assistant to search and share in-device photos. meChat's ability to provide a stand-alone solution utilizing on-device intelligence enables it in searching highly relevant photos with low perceived latency and energy consumption while preserving user privacy.

Next, in *"An Embodied Cognitive Assistant for Visualizing and Analyzing Exoplanet Data,"* Jeffrey Kephart *et al.* share a rather unorthodox application of a chatbot. They demonstrated an embodied cognitive agent that is capable of assisting astrophysicists to visualize and analyze exoplanet data via natural interactions with a combination of multimodal inputs via speech and gestures. They aimed to realize the vision of symbiotic cognitive computing where software agents coinhabit a physical space with people and use their understanding within a specific domain to act as valuable collaborators on cognitive tasks. This reduces the mental expenditure in data visualization while channeling the focus on more creative and exploratory pursuits. The paper also described some key functionalities such as Deixis via simultaneous speech and gesture, progressive querying and iterative refinement, minimizing reliance on attention word, seeking clarification, symbolic model discovery, and explainable self-programming using an AI planner for derivative tasks which are essential for naturalness of interaction.

Finally in *"Bots Acting Like Humans: Understanding and Preventing Harm,"* Florian Daniel *et al.* discuss the social implications of dealing with bots. They propose a foundational framework for bot ethics by curating a taxonomy of bot harms, exemplified with concrete examples of bot failures (causing harm). The key contribution is to motivate the need of ethical guidelines for virtual bots by creating a common understanding of the types of harm and abuse induced in the event of bot failures given its strong growth in presence in public and private communications. They present how bots can be intentionally or unintentionally abused, leading to psychological, legal, economic, social, and democratic harms. The paper also follows up with approaches in preventing such abuses such as bot banning and explicit declaration of bot usage to end users. Content and behavior analysis techniques using crowdsourcing, NLP, and ML to mitigate some of the technical and ethical challenges are also discussed.

All in all, a chatbot is no longer an incomprehensible rule-based conversational interface that spits the stilted "Sorry, I didn't understand that" or "Sorry, I don't have an answer for that." Today, it has evolved into smarter personal digital assistants powered by a paradigmatic set of behind-the-scene AI cognitive services equipped with semantic, cognitive, and perceptual computing techniques, along with ML, deep learning, and NLP that encompasses natural language understanding and natural language generation to provide better context for richer and more natural interactions.[8] Enabling "Turing-like" interactions with technology has always been a long standing promise. Such levels of intelligence in communication like recognizing the user's needs based on prior knowledge and communications and subsequently extracting key intent, analyzing, interpreting in the right context, and personalizing responses with a right situational awareness will ultimately propel us to the realm of humanoid augmented reality assistants (http://bit.ly/Humanlike-Chatbots).

ACKNOWLEDGMENTS

■ REFERENCES

1. S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Upper Saddle River, NJ, USA: Prentice Hall, 2009.

2. R. Daft, and R. Lengel, "Organizational information requirements, media richness and structural design," *Manage. Sci.*, vol. 23, no. 20, pp. 554–571, 1986, doi: 10.1287/mnsc.32.5.554.

3. R. Lengel and R. L. Daft, "The selection of communication media as an executive skill," *Acad.*

Manage. Executive, vol. 2, pp. 225–232, Aug. 1989, doi:10.5465/ame.1988.4277259.

4. M. Kalantari, "Consumers' adoption of wearable technologies: Literature review, synthesis, and future research agenda," *Int. J. Technol. Marketing*, vol. 12, no. 3, p. 274, 2017, doi: 10.1504/ijtmkt.2017.089665.

5. A. Sheth, "Smart data - How you and I will exploit big data for personalized digital health and many other activities," in *Proc. IEEE Int. Conf. Big Data*, 2014, doi: 10.1109/bigdata.2014.7004204.

6. IBM Watson Products and Services, 2018. [Online]. Available: https://www.ibm.com/watson/products-services/

7. Cognitive Services - Microsoft Azure, 2018. [Online]. Available: https://azure.microsoft.com/en-us/services/cognitive-services/

8. A. Sheth, "Internet of Things to smart IoT through semantic, cognitive, and perceptual computing," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 108–112, Mar./Apr. 2016, doi: 10.1109/mis.2016.34.

9. U. Jaimini, K. Thirunarayan, M. Kalra, R. Venkataraman, D. Kadariya, and A. Sheth, "How is my child's asthma?," Digital Phenotype and Actionable Insights for Pediatric Asthma, 2018.

10. N. Nag, V. Pandey, P. J. Putzel, H. Bhimaraju, S. Krishnan, and R. C. Jain, "Cross-modal health state estimation," in *Proc. 2018 ACM Multi. Conf.*, Seoul, Republic of Korea, 2018, https://doi.org/10.1145/3240508.3241913

11. K. Fitzpatrick, A. Darcy, and M. Vierhile,. "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial," *JMIR Mental Health*, vol. 4, no. 2, p. e19, 2017, doi: 10.2196/mental.7785

12. A. Sheth, U. Jaimini, K. Thirunarayan, and T. Banerjee, "Augmented personalized health: How smart data with IoTs and AI is about to change healthcare," in *Proc. IEEE 3rd Int. Forum Res. Technologies Soc. Industry*, Sep. 11–13, 2017, pp. 1–6.

Amit Sheth is the LexisNexis Ohio Eminent Scholar and the Executive Director with Ohio Center of Excellence, Knowledge-enabled Computing, Wright State University, Dayton, OH, USA. He is a fellow of IEEE, AAAI, and AAAS. Contact him at amit@knoesis.org.

Hong Yung (Joey) Yip is currently working toward the Ph.D. degree at Kno.e.sis Center, Wright State University, Dayton, OH, USA. His research interests include conversational AI, applied machine learning, and semantic web including development of chatbots for medical applications. Contact him at joey@knoesis.org.

Arun Iyengar does research, development, and consulting in cloud computing and artificial intelligence at IBM's T.J. Watson Research Center. He is an IEEE Fellow. He was the recipient of several Best Paper awards, the IFIP Silver Core Award, and been named an IBM Master Inventor multiple times. Contact him at aruni@us.ibm.com.

Paul Tepper is a Computational Linguist, Product Manager, and Software Engineer with a background in artificial intelligence, linguistics & cognitive science. He is currently the Head of the Cognitive Innovation Group AI Lab and Product Manager for AI & Machine Learning at Nuance Communications, Enterprise Division. Contact him at paul.tepper@nuance.com.