

## Deep Learning for Domain Scaling of Conversational Agents

Intelligent Agents/chat-bots have become a hot topic in industry. *Amazon, Apple, Google, Facebook* and *Microsoft* have all invested heavily in the area. Many start-ups work on different perspective of the space as well, ranging from language understanding techniques to solutions for specific tasks (e.g., appointment scheduling). However, it is still very costly to introduce a new experience to an agent/bot. A major issue here is that language understanding and conversation management modeling are often performed in a domain-specific fashion – either with data-driven statistical modeling or with semantic grammar authoring – the former requires a large amount of labeled training data; the latter needs the combined expertise in linguistics and domain knowledge. In this talk, we formulate the domain scaling as a training data demand-supply problem, and introduce some preliminary investigations and experiment results on this problem.

Ye-Yi Wang is currently a Partner Research Manager at *Bing, Microsoft*. He obtained his Ph.D. in Language technology from *Carnegie Mellon University*, under the supervision of Prof. Waibel, in 1998. He joined the speech technology group in *Microsoft Research* right after his graduate study, working on language modeling, spoken language understanding, conversational systems, query understanding, and information retrieval. After a 12-year career in *Microsoft Research*, he joined *Bing* in 2010. He has been working on query understanding, web search relevance, entity linking/ranking (mapping queries to entities in a knowledge graph), and the machine learning platform for query processing. The platform has later become the foundation for language understanding and dialog management of Cortana. He has led his team to deliver the Cortana language understanding and dialog models, as well as bring the intelligence to *Microsoft Office* by building the relevance platform and solutions in *Microsoft Exchange* service. He is very keen on addressing the domain scaling and third party extensibility problem for Cortana and other conversational agent/chat bots.