

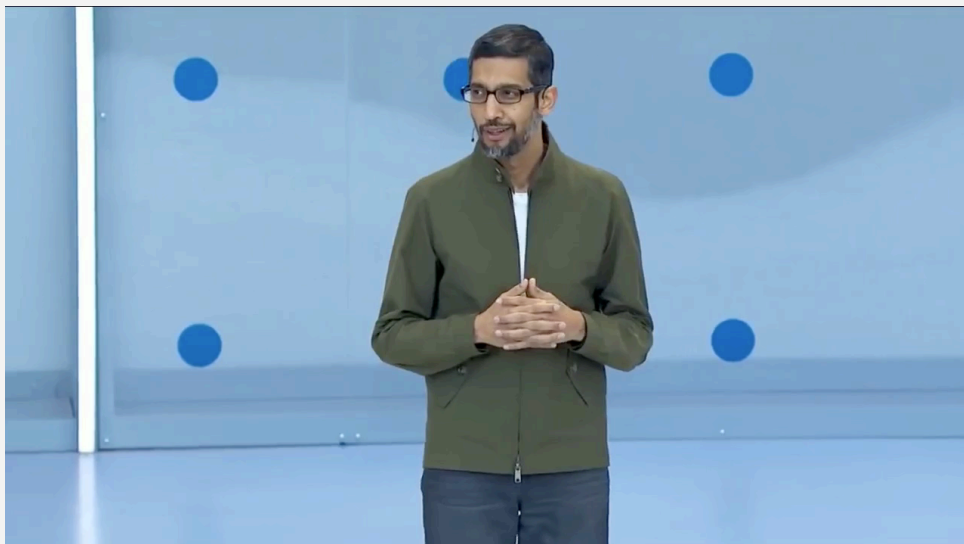


RESEARCH TOPICS

NGUYEN THI THU TRANG
TRANGNTT@SOICT.HUST.EDU.VN

1

GOOGLE DUPLEX: AI ASSISTANT CALLS LOCAL BUSINESSES TO MAKE APPOINTMENTS (2017)



Text-To-Speech

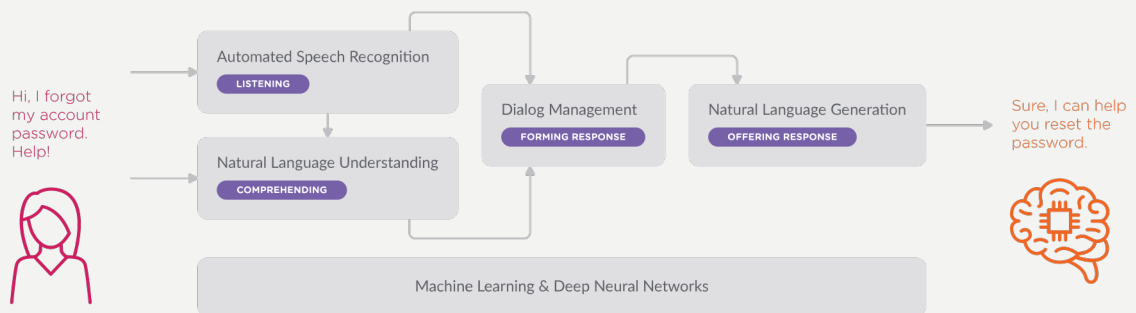
Natural Language Understanding

Automatic Speech Recognition

2

CONVERSATIONAL AI

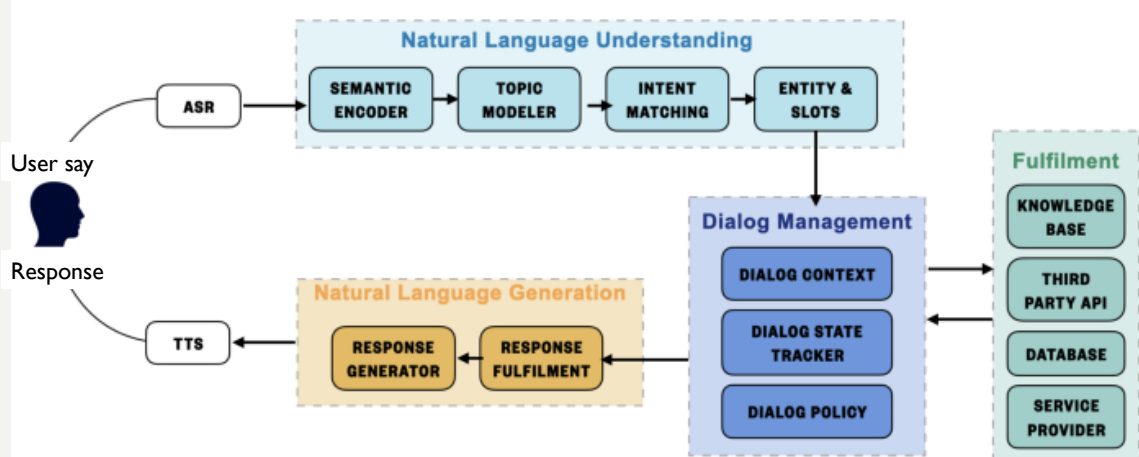
- A set of technologies behind automated messaging and speech-enabled applications that offer human-like interactions between computers and humans



Automatic Speech Recognition (ASR), Text To Speech (TTS), Natural Language Processing (NLP), Advanced Dialog management, and Machine Learning (ML) to understand, react and learn from every interaction

3

CONVERSATIONAL AI ARCHITECTURE



4

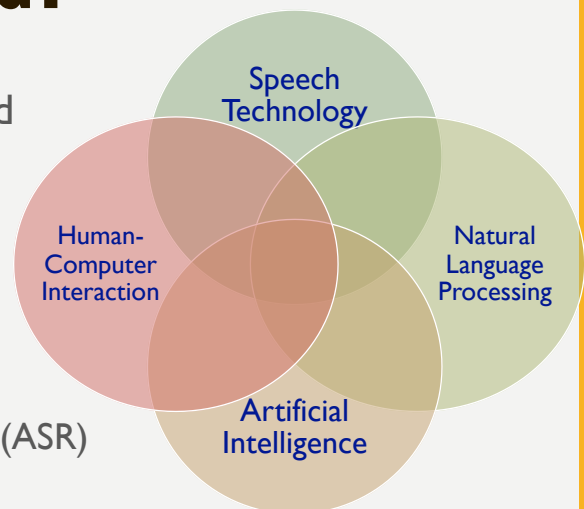
TEXT PROCESSING

- Text Pre-processing
 - Language Identification
 - Spelling Correction
 - Text Normalization
 - Spam/Noise/Hate Speech Detection
- Conversational AI
 - Multi-message Chat Management
 - Dialog Management
 - Multi-Intent Classification
 - Context-aware Intent Classification
- Text Sentiment Analysis

5

SPEECH TECHNOLOGY

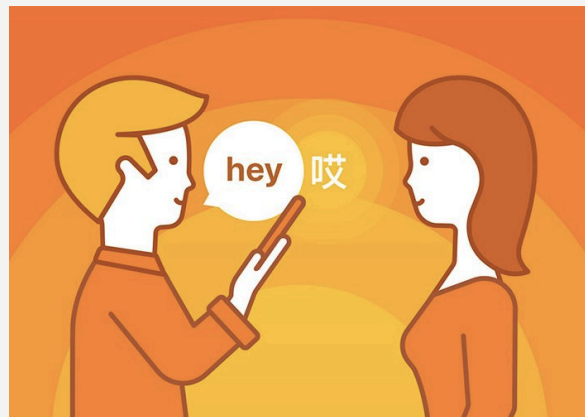
- Recognition, understanding and generation of human speech
- Main challenges
 - Speech Modeling
 - Text-To-Speech (TTS)
 - Automatic Speech Recognition (ASR)
 - Spoken Dialogue Systems



6

APPLICATIONS OF SPEECH TECHNOLOGY

- Virtual Assistant
- Speech Translation
- User Authentication
- Call Center with Virtual Agents
- Screen Readers
- ...



7

RESEARCH TOPICS ON SPEECH TECHNOLOGY

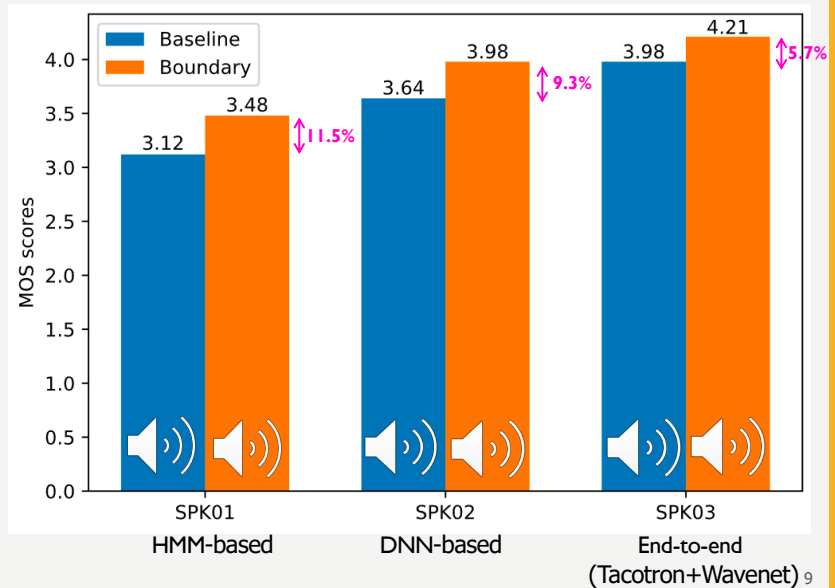
- Speech Synthesis (Text-To-Speech)
 - Text Normalization
 - Speaker Adaptation
 - Expressive Speech Synthesis
 - State-of-the-art Synthesis Techniques
- Automatic Speech Recognition (ASR)
 - Speech-To-Text
 - Speaker Recognition (Identification/Verification)
 - Speech Sentiment Analysis

8

SPEECH SYNTHESIS

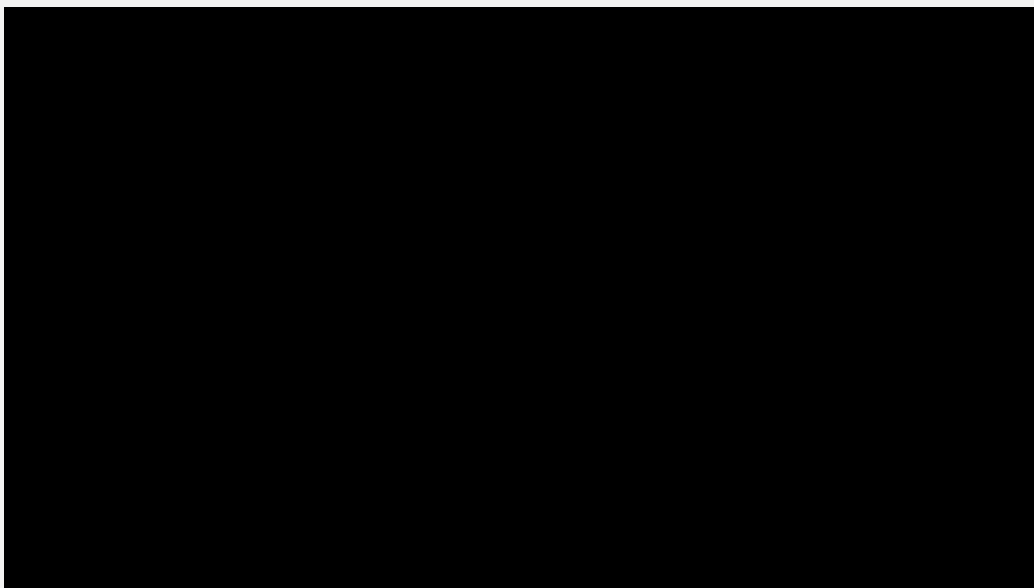
Nguyen Thi Thu Trang,
 Nguyen Hoang Ky, Albert
 Rilliard, Christophe
 d'Alessandro. **Prosodic
 Boundary Prediction
 Model for Vietnamese
 Text-To-Speech**. In
 Proceeding of 22th
 International Speech
 Communication Association
 (Interspeech 2021). Brno,
 Czech Republic. September
 2021.

- **SPK01**: Female Northern
- **SPK02**: Male Northern
- **SPK03**: Female Northern



9

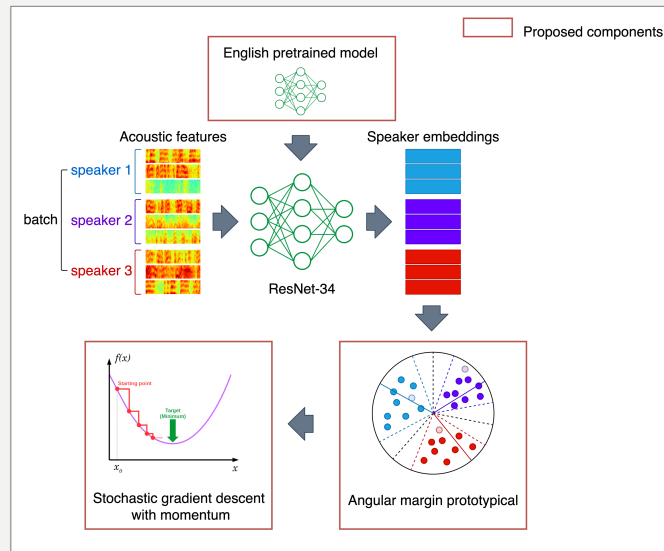
VOICE BIOMETRICS FOR CUSTOMER IDENTIFICATION (NUANCE, US)



10

SPEAKER VERIFICATION

- Vi Thanh Dat, Pham Viet Thanh, Nguyen Thi Thu Trang. **Deep Speaker Verification Model for Low-Resource Languages and Vietnamese Dataset**. In Proceedings of 35th Pacific Asia Conference on Language, Information and Computation (PACLIC 2021). Shanghai, China. November 2021.



11

RESEARCH & DEVELOPMENT

- Research
 - Proposal of new/improved model / architecture
 - Corpus Building & Experiments
 - Publications
- R&D
 - Develop models applied in the real-life projects
 - Hybrid models: Heuristics, EA, ML, DL
 - Tuning / Optimizing with real-life constraints: Noisy, Real-time, Robust

12