



# WORKSHOP ON AUTOMATIC SPEECH RECOGNITION (ASR) ON EMERGENCY MEDICINE



NEURALMECHANICS.NET

-  @NeuralMechanics
-  @nmi-ph
-  @nmi\_ph

Prepared by: Jay Garcia

Presented by: Isabel Saludares

-  @mismabs
-  @mismabs

# COMMUNICATION

## Typical communication episode

*S (speaker) wants to convey P (proposition) to H (hearer) using W (words in a formal or natural language)*

### Speaker

**Intention:** S wants H to believe P

**Generation:** S chooses words W

**Synthesis:** S utters words W

### Hearer

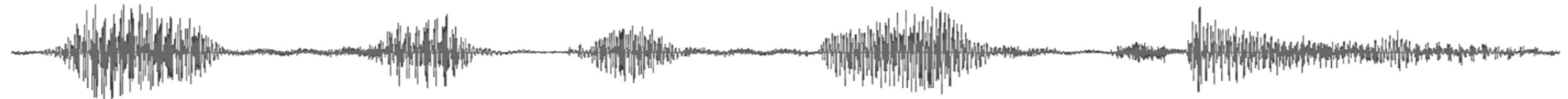
**Perception:** H perceives words W"  
(ideally  $W'' = W$ )

**Analysis:** H infers possible meanings  
P1,P2,...,Pn for W"

**Disambiguation:** H infers that S intended  
to convey Pi (ideally  $P_i = P$ )

**Incorporation:** H decides to believe or  
disbelieve Pi

Audio



## Automatic Speech Recognition

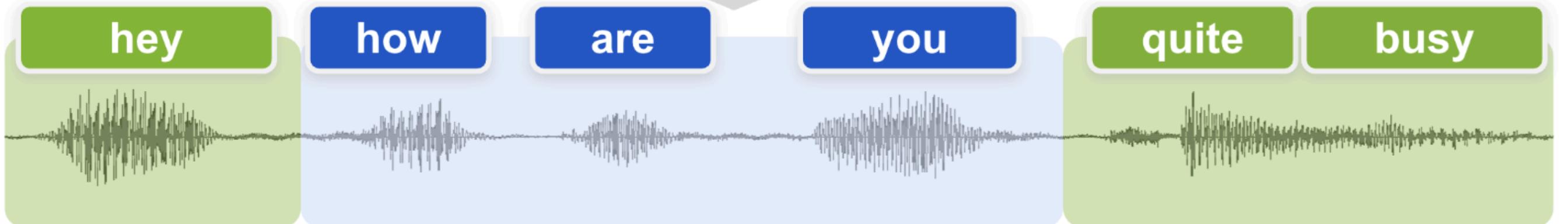
Transcription

hey how are you quite busy



## Speaker Diarization

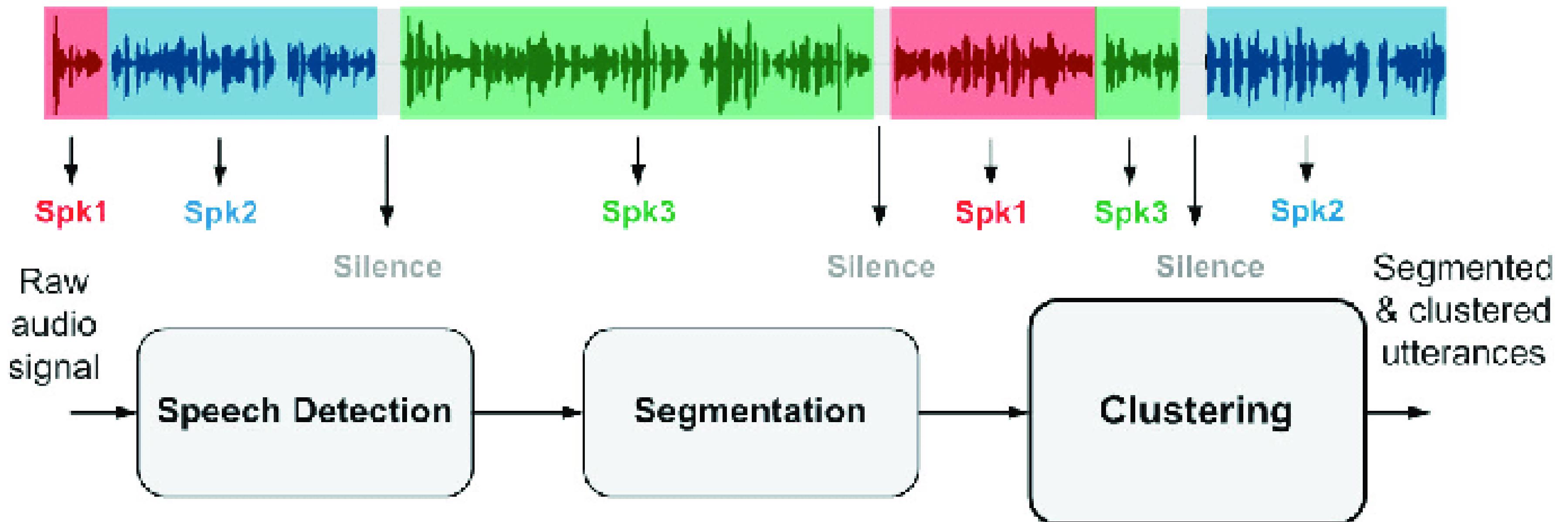
Speaker Labels



Speaker 1

Speaker 2

Speaker 1



# NATURAL LANGUAGE PROCESSING



MonkeyLearn

Very **intuitive platform**, I'll **definitely recommend** it.

The **chat support** is **excellent**, really **fast** in their replies  
and very **helpful**.

**Usability**

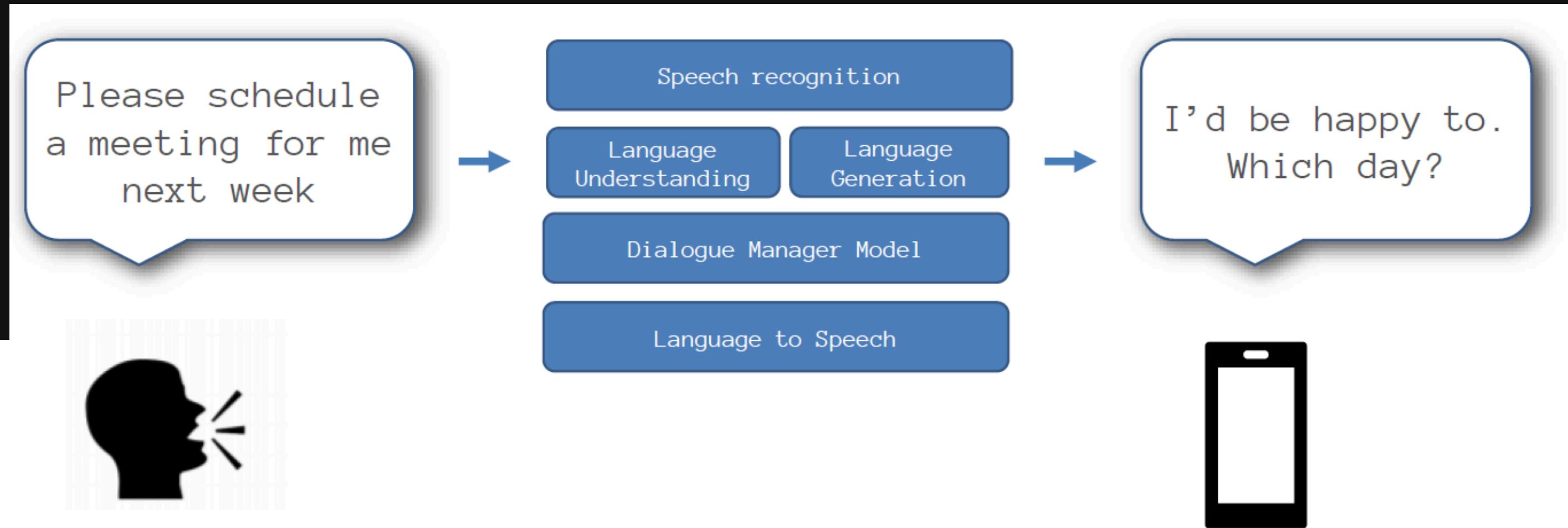
**Positive**

**Customer Support**

## understanding

*Taking some spoken/typed sentence and working out what it means*

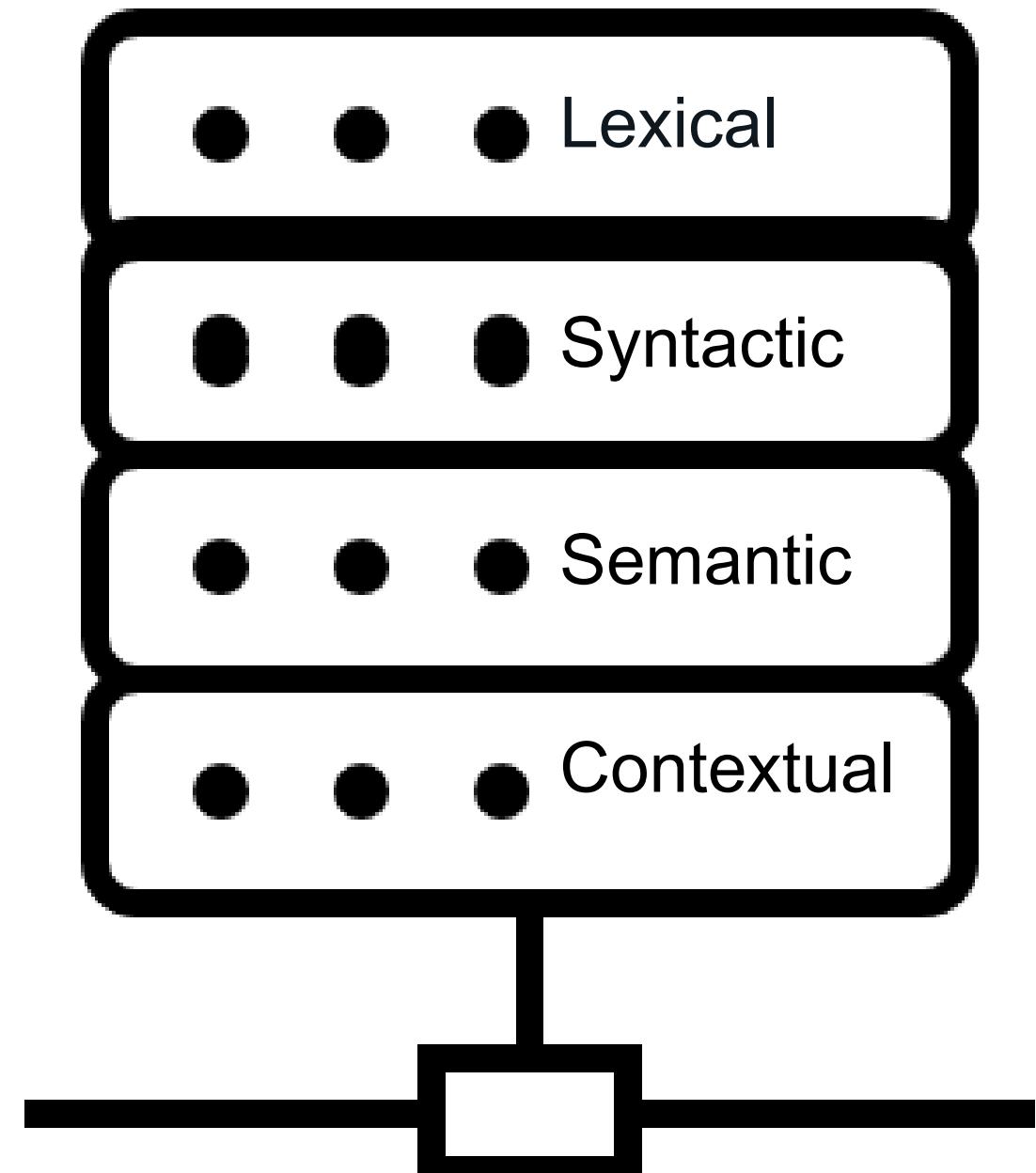
# NATURAL LANGUAGE PROCESSING



## generation

*Taking some formal representation of what you want to say and working out a way to express it in a natural (human) language (e.g., English)*

# NLP Layers



Basic properties of words → Spell check, NER

Order and structure of words → Grammar check

Meaning of words → WordNet, etc.

Overall meaning of text → Topic modeling, sentiment analysis





# Generative AI in Healthcare with Nuance and Epic at Microsoft Inspire 2023



Share



The screenshot shows the Microsoft HealthVault interface. On the left, there's a sidebar with navigation links like 'In Boxlist', 'My Messages', 'Results', 'My Medical Actions', 'Re Request', 'My Open Charts', 'Chart Comparison', 'E Visits', 'Message List', 'Attached to Coming Up...', 'Follow-up', 'Search', 'Send Messages', and 'Completed Work'. The main area has tabs for 'My Medical Actions' and 'New Grandparent'. A large red play button is overlaid on the interface. The 'New Grandparent' tab is active, showing a message from 'George Adams, Jr.' to 'Dr. Walker'. The message content is:

**Generated Draft Reply**

Subject: New grandparent

Congratulations on becoming a new grandparent! I would like to hear about your extended family. In particular, I would like to know about your mother's side of the family to maintain good hygiene or advice, such as visiting your friends frequently and avoiding close contact if you're sick. If you have specific concerns about your mother's side of your family health, I would recommend scheduling an appointment to discuss this further.

Please don't hesitate to reach out to me. Please any other questions or concerns.

Dr. Walker

George Adams, Jr.

**New grandparent**

To: Dr. Walker

Subject: New grandparent

Hi Dr. Walker, I thought I'd tell you that I'm a new grandparent! That's adorable but also a frequently asked question. Do I need to be concerned that being around her parents will hurt my work value?

**Enclosed Messages**

From: George Adams, Jr. To: Dr. Walker

Subject: New grandparent

Date: Composed: 8/10/2023 4:03 AM

Watch on YouTube

# HEALTHCARE AREAS FOR ASR



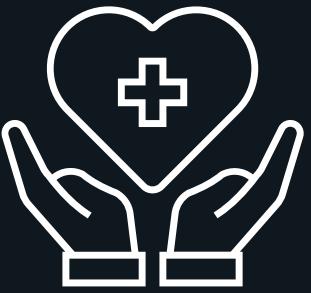
**5 areas in healthcare  
where ASR and  
analytics can create  
value**



diagnosis, detection, and  
tracking of disease



medical transcription and  
digital transcribing



enhance adherence to plan  
of care among patients



improve quality of lives for  
senior and the disabled



customer communication  
insights for payers and  
providers

# MEDICAL TRANSCRIPTION

- Transforms physicians' or patients' **voice input into consistent written reports, appointment summaries, treatment plans, mood journal entries, symptom summaries**, etc.
- Then, this data is uploaded to EHR, a patient or RPM app, or other target software.
- **Reducing time spent in documenting** patient visits into electronic health record (EHR) systems
- **Language support** for Tagalog, English, and Taglish variants

# DIAGNOSIS, DETECTION OF DISEASES

- Use of **vocal biomarkers to detect and diagnose diseases** ahead of time
  - Vocal biomarkers → patient's voice, speech patterns and breathing rhythm hold clues to underlying conditions
- Possible **early detection of Alzheimer's and Dementia by using a patient's speech patterns**

# IMPROVE PATIENT'S QUALITY OF LIFE

- Doctors use **speech recognition-powered virtual assistants** to schedule appointments, tests, and diagnostic procedures, create and retrieve health records on the go.
- Patients with speech or hearing difficulties can be improved by converting speech to text or vice versa

# CUSTOMER COMMUNICATION AND ANALYTICS

- **Voice analytics** helps provide inputs to such programs by:
  - a. Making highly **accurate call transcripts** available by using appropriate speech-to-text technologies
  - b. Applying **text mining** (keyword spotting, etc.) and natural language processing methods or NLP (sentiment analysis, entity extraction, summarization, etc.) to provide intelligence and insights.
- Command center quality audits to assess agent/patient behavior and conversation flow during a call

# IMPROVE PATIENT ADHERENCE

- Routinely asking questions regarding a patient's condition and offering reminders
- **AI assistants** that can provide clinically approved advice and recommendations at home, etc.

# SPEECH RECOGNITION



**Amazon  
Transcribe**



Azure  
Cognitive Services



Azure  
Speech to Text API



Google Cloud  
Speech API



Dragon APIs



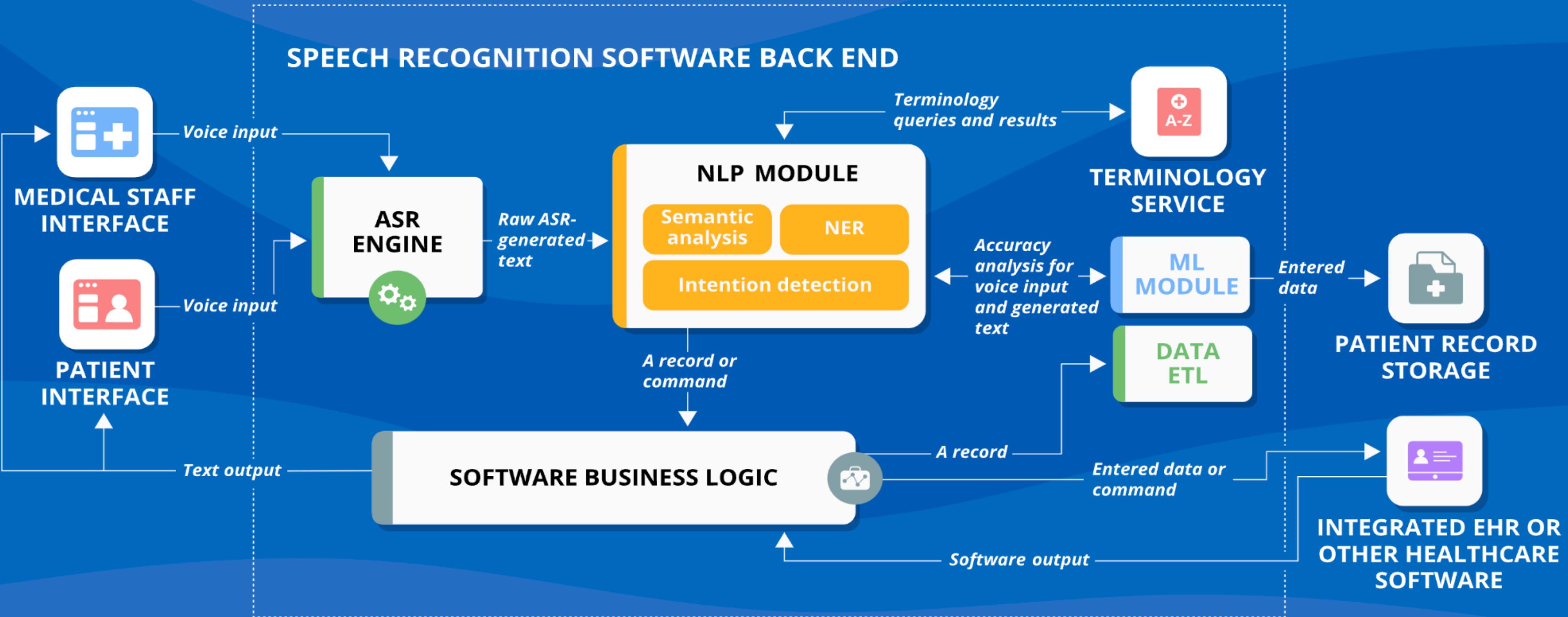
IBM Watson



Whisper

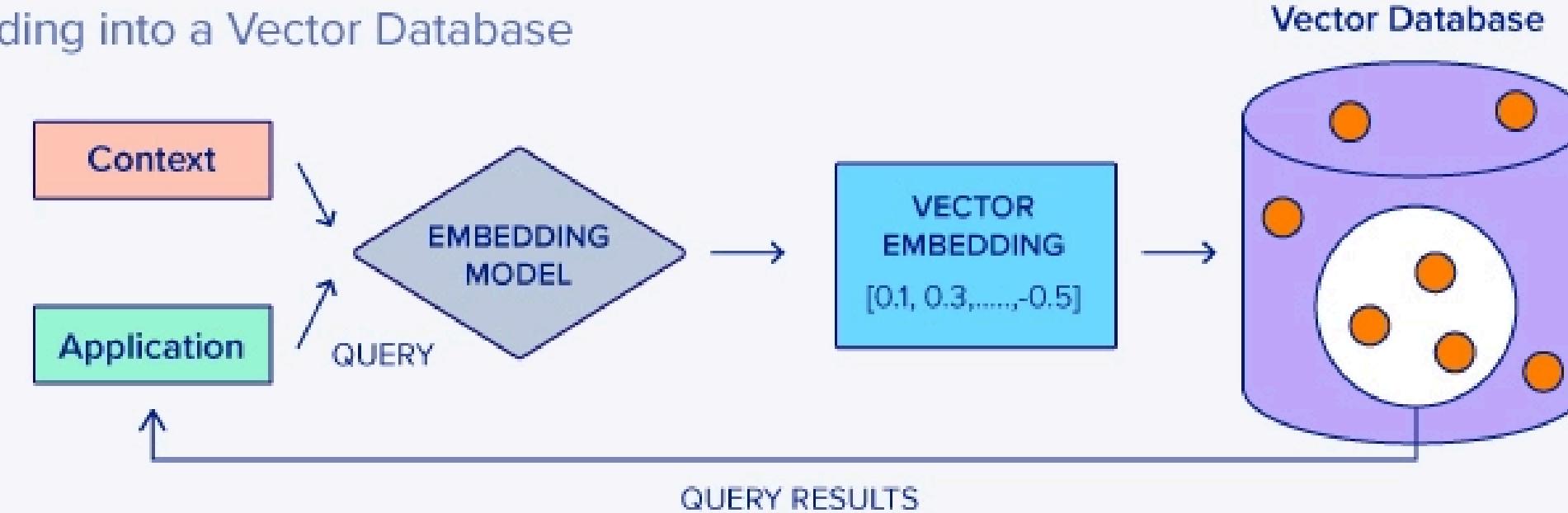


# CASE STUDIES (DEMONSTRATION)



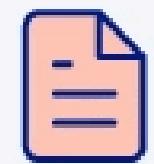
# GENERATIVE AI ON AUDIO TRANSCRIPTS

Embedding into a Vector Database



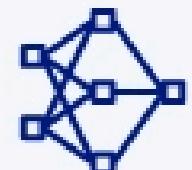
Vector Database

Document



TAKES IN  
DATA

AI Model



OUTPUTS  
VECTOR  
EMBEDDING

Vector Database



Answer



Retrieving a Vector Embedding

TRANSLATE VECTOR  
TO NATURAL LANGUAGE



Query

# INTEGRATION

- An **automatic speech recognition (ASR) engine** transforms **voice input into text**. Then, a **natural language processing (NLP)** module helps interpret the voice data by using:
  - **Semantic analysis** that helps adjust the ASR-generated text based on the context and make it cohesive
  - **Named entity recognition (NER)** technology that detects certain entities within the text (e.g., a person, a health organization, a condition) and checks the text against publicly available knowledge bases (e.g., Unified Medical Language System) to generate a health record.
- Intention detection that identifies voice commands and sends them to the software business logic for execution.

# INTEGRATION

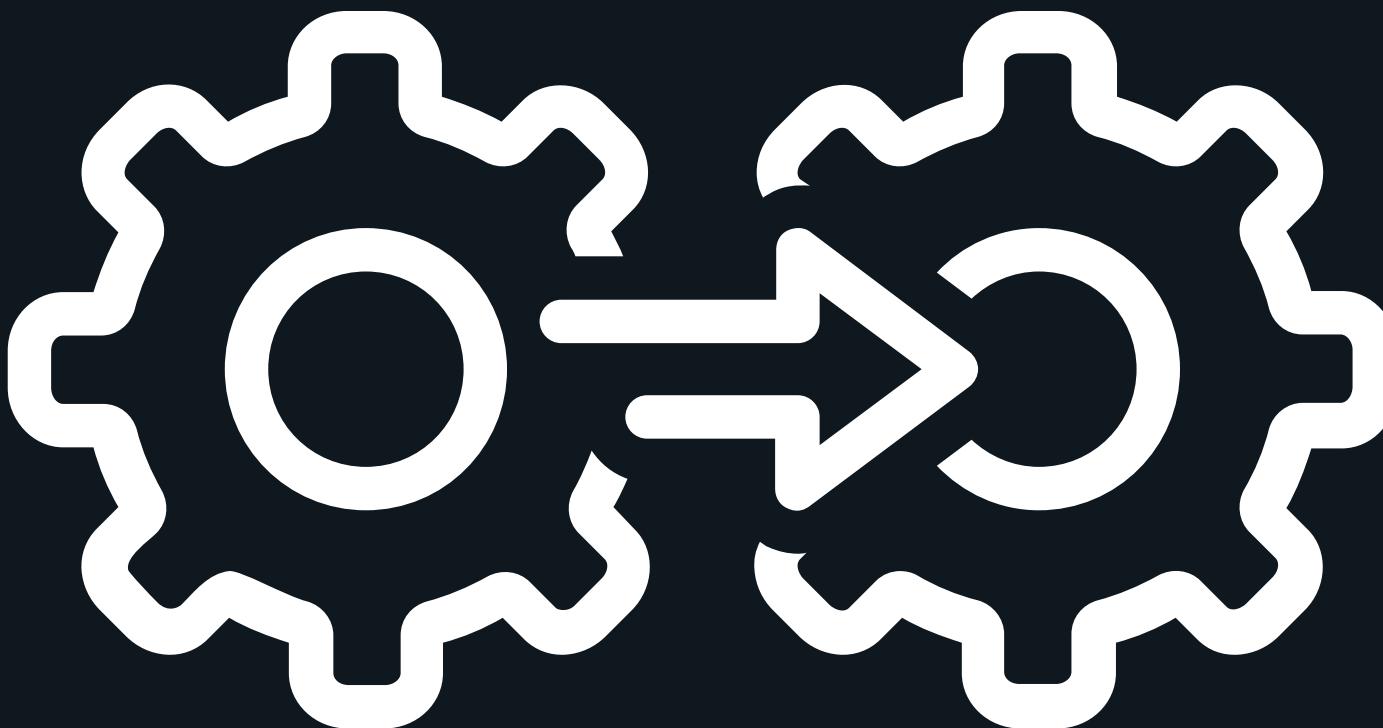
- The **NLP module** is connected to a terminology service featuring various medical terms, popular abbreviations, etc.
- The **voice recognition software** may be powered by an additional machine learning module to improve speech recognition quality or adjust to specific speech patterns and accents.

# ASSESSING YOUR CURRENT SYSTEM

- Estimated **volume** (ie. number of patients, incoming phone calls)
- Describe your **patient information journey**:
  - patient conversation --> transcription --> archiving --> integration to other hospital systems
- What **EHR system** are you using? Other software for your patient records management?

# How do we integrate our Speech Recognition system within the hospital system?

*via API service*  
*separate interface*  
*separate hardware*



# THANK YOU!

## PLEASE VISIT OUR BOOTH FOR MORE DETAILS



@NeuralMechanics



@nmi-ph



@nmi\_ph



NMI Social Pages