

Sobre el uso de IA en la curaduría de las comunicaciones de voz de las misiones Apolo de la NASA

Iván López-Espejo

I Encuentro Iberoamericano sobre Archivos e Inteligencia Artificial
Cooperación para la Preservación de la Memoria en la Era Digital

iloes@ugr.es

29 de octubre de 2024



Secretaría General
Iberoamericana
Secretaria-Geral
Ibero-Americana



RED DE ARCHIVOS
DIPLOMÁTICOS
IBEROAMERICANOS

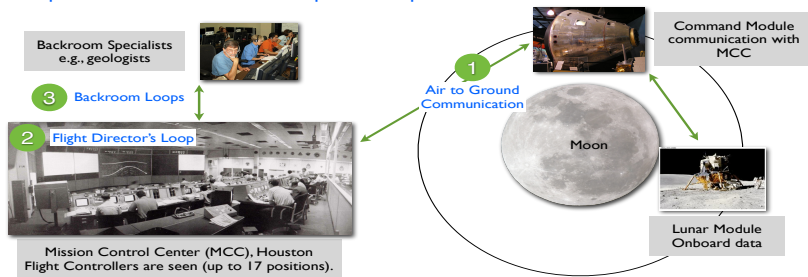


- 1 Introducción
- 2 Proyecto AGILE-KWS
- 3 Detección de Palabras Clave (KWS)
- 4 Identificación de Tema a partir de KWS
- 5 Identificación de Tema de Extremo a Extremo
- 6 Bibliografía

Introducción

- Las **misiones Apolo de la NASA** destacan como uno de los mayores logros de la humanidad
- La NASA **grabó en cintas analógicas** todas las conversaciones entre los astronautas, los especialistas de control de misión y el personal de apoyo que hicieron que estas misiones fueran un éxito

Complex Interaction results in multiple Audio/Speech data sources



A. Sangwan *et al.*, "'Houston, We have a solution': Using NASA Apollo Program to advance Speech and Language Processing Technology," en *Proc. de INTERSPEECH 2013*

Introducción

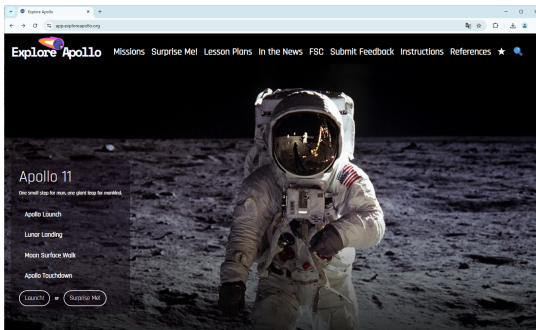
- Las **misiones Apolo de la NASA** destacan como uno de los mayores logros de la humanidad
- La NASA **grabó en cintas analógicas** todas las conversaciones entre los astronautas, los especialistas de control de misión y el personal de apoyo que hicieron que estas misiones fueran un éxito



Introducción

- Subvención de la National Science Foundation (**NSF**) para convertir el vasto archivo de cintas de audio en **“Explore Apollo”** (exploreapollo.org)
- “Explore Apollo” es un sitio web de educación e investigación que brinda **acceso público a los materiales de voz**

Prof. John H. L. Hansen



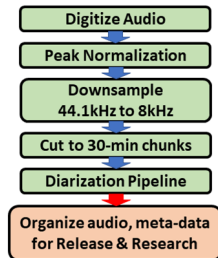
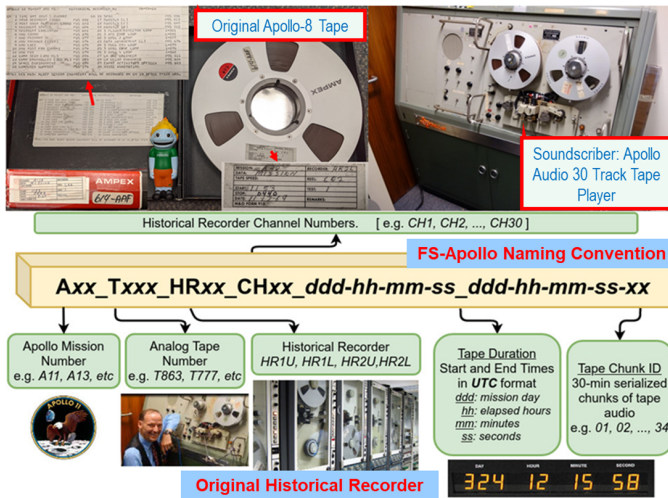
Introducción

- **CRSS** (Center for Robust Speech Systems) está haciendo posible la conservación de una valiosa pieza histórica
- Ardua tarea consistente en la **digitalización de cientos de cintas analógicas** de 14 horas de duración cada una
- Producción de un corpus de habla real, único en su clase, de **decenas de miles de horas**
 - ▶ Apollo 11
 - ▶ Apollo 13
 - ▶ Gemini 8
 - ▶ ...

Prof. John H. L. Hansen



Introducción



A. Joglekar et al., "Fearless Steps: Advancements in Speech Technology and Corpus Development for Naturalistic Audio," en 2023 NASA HRP IWS

Introducción

- La disponibilidad de un corpus de habla real de estas características ofrece la oportunidad de un **salto de rendimiento para las tecnologías del habla**
- **The Fearless Steps Challenge**
 - ▶ Detección de actividad de voz
 - ▶ Identificación de locutor
 - ▶ Diarización de locutor
 - ▶ Reconocimiento automático del habla (ASR)
 - ▶ Análisis conversacional
- app.exploreapollo.org emplea algoritmos de diarización y ASR desarrollados por CRSS para facilitar la consulta de los materiales de voz y proporcionar una nueva perspectiva sobre las misiones Apolo de la NASA



Introducción

- Aún en proceso de recuperación del resto del audio de las misiones Apolo
- ¡En total, se espera recuperar **más de 150.000 horas de audio!** → **El corpus de habla natural basado en comunicaciones reales más grande existente**

Disponer de **mecanismos de indexación de metadatos** que garanticen una consulta eficaz de los materiales de voz resulta imperativo

A Glant LEap for KeyWord Spotting (**AGILE-KWS**)

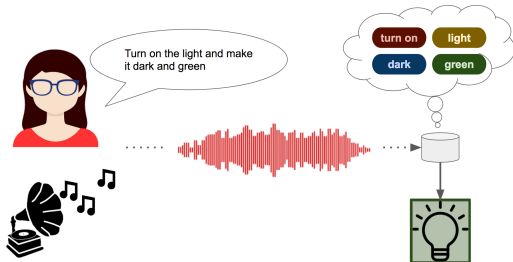
- **Clasificación por tema** de los diferentes fragmentos sonoros
 - ▶ Conversaciones personales
 - ▶ Problemas técnicos
 - ▶ ...
- Anotación automática ($j > 150k$ horas!) de las grabaciones
 - ▶ **Detección de palabras clave** (**KWS**) vs. ASR

Marie Skłodowska-Curie Action, Global Postdoctoral Fellowship
(2022-2025)



Detección de Palabras Clave (KWS)

- La **detección de palabras clave (KWS)** trata de la identificación de palabras clave en señales de audio
- Las palabras clave se preseleccionan en base al conocimiento del dominio específico



T. Bluche *et al.*, "Small-Footprint Open-Vocabulary Keyword Spotting with Quantized LSTM Networks," arXiv, 2020

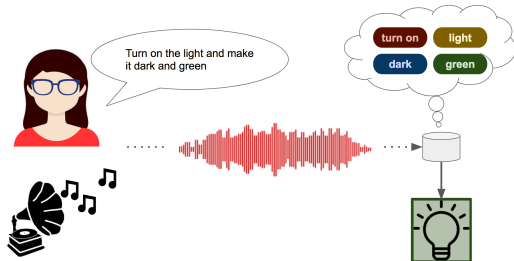
• Aplicaciones

- 1 Activación de dispositivos
- 2 Identificación de comandos
- 3 Detección de sentimientos
- 4 Recuperación de información
- 5 Identificación de tema

Detección de Palabras Clave (KWS)

¿Por qué necesitamos KWS?

- La extracción de información de alto nivel (p. ej., tema) a partir de audio natural es complicada debido al **ruido y otros factores acústicos**
- **KWS simplifica el proceso** decodificando el espacio acústico y filtrando los segmentos relevantes para la comprensión conversacional

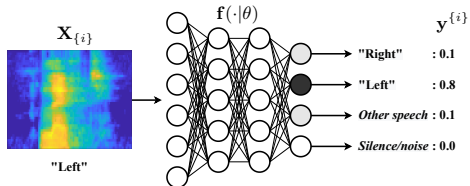
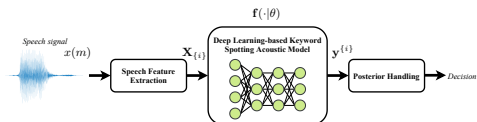


T. Bluche *et al.*, "Small-Footprint Open-Vocabulary Keyword Spotting with Quantized LSTM Networks," arXiv, 2020

• Aplicaciones

- 1 Activación de dispositivos
- 2 Identificación de comandos
- 3 Detección de sentimientos
- 4 Recuperación de información
- 5 Identificación de tema

Detección de Palabras Clave (KWS)



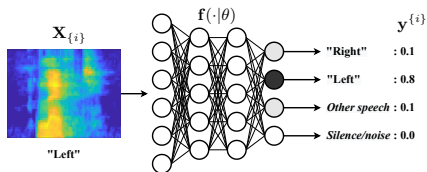
I. López-Espejo *et al.*, "Deep Spoken Keyword Spotting: An Overview," IEEE Access, 2022

- 1 $f(\cdot|\theta)$ se implementa mediante una **red neuronal profunda**

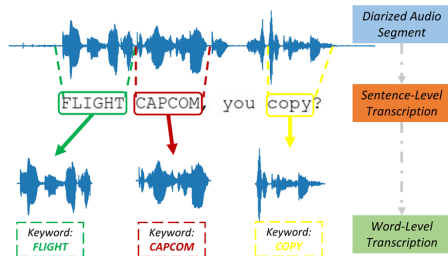
- AGILE-KWS perseguía el desarrollo de **KWS robusto a distorsiones acústicas**

- 1 Ruido de canal
- 2 Ruido de sistemas
- 3 Atenuación del ancho de banda de la señal
- 4 Ruido de transmisión
- 5 Ruido cósmico
- 6 Ruido estático de las cintas analógicas
- 7 Ruido debido al envejecimiento de las cintas
- 8 Ruido ambiente
- 9 Reverberación
- 10 Fricción del traje espacial al caminar en la Luna
- 11 Tonos Quindar

Detección de Palabras Clave (KWS)



I. López-Espejo et al., "Deep Spoken Keyword Spotting: An Overview," IEEE Access, 2022

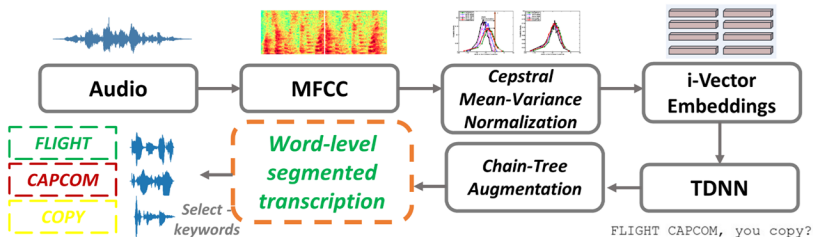


A. Joglekar et al., "Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting," en 2024 NASA HRP IWS

- Sólo hay transcripciones manuales a nivel de frase para 125 horas de audio
- ¡Necesitamos **segmentaciones a nivel de palabra** para desarrollar sistemas de KWS!
- **Solución:** Obtención de marcas de tiempo a nivel de palabra mediante **alineamiento forzado**

Detección de Palabras Clave (KWS)

Solución propuesta de alineamiento forzado



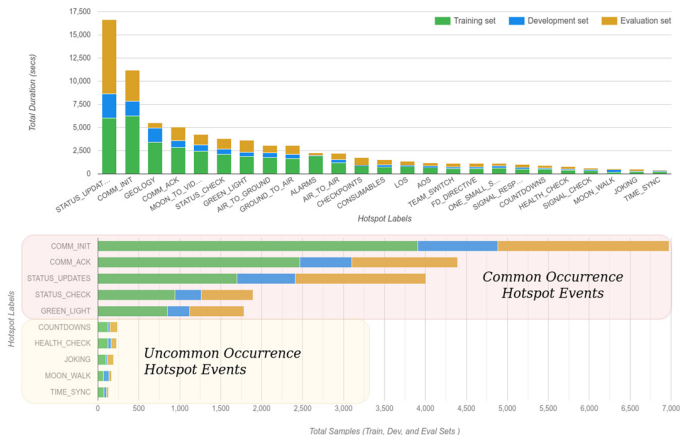
Cuadro: Precisión (%) de palabra clave tras segmentación mediante alineamiento forzado

| Sistema | MFA | FeaRLESS – FA | TRILL – FA | Propuesta FA |
|---------------|-----|---------------|------------|--------------|
| Precisión (%) | 93 | 89 | 94 | 99 |

A. Joglekar et al., "Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting," en 2024 NASA HRP IWS

Identificación de Tema a partir de KWS

- 25 temas de conversación identificados
- Temas de conversación anotados manualmente para 115 horas de audio
- En términos de duración, los temas se encuentran considerablemente desbalanceados



A. Joglekar *et al.*, "Fearless Steps APOLLO: Challenges in keyword spotting and topic detection for naturalistic audio streams,"
The Journal of the Acoustical Society of America, 2023

Identificación de Tema a partir de KWS

Ciertas **palabras clave** son identificativas de determinados **temas de conversación**

Illustration: Apollo 11



CAPCOM: FLIGHT CAPCOM, you copy? [conv_label: **COMM_INIT**]

Speaker: Charlie Duke; Mission Role: Capsule Communicator; SpeakerID: CAPCOM1



FD: Go ahead FLIGHT. [conv_label: **COMM_ACK**]

Speaker: Gene Kranz; Mission Role: Flight Director; SpeakerID: FD1

Illustration: Apollo 13



CAPCOM: This is Houston.
Say again, please. [conv_label: **COMM_ACK**]

Speaker: Jack Lousma; Mission Role: Capsule Communicator; SpeakerID: CAPCOM13



LOVELL: Ah, Houston,
we've had a problem [conv_label: **ALARMS**]

Speaker: Jim Lovell; Mission Role: Flight Commander (Astronaut); SpeakerID: LOVELL

A. Joglekar et al., "Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting," en 2024 NASA HRP IWS

Identificación de Tema a partir de KWS

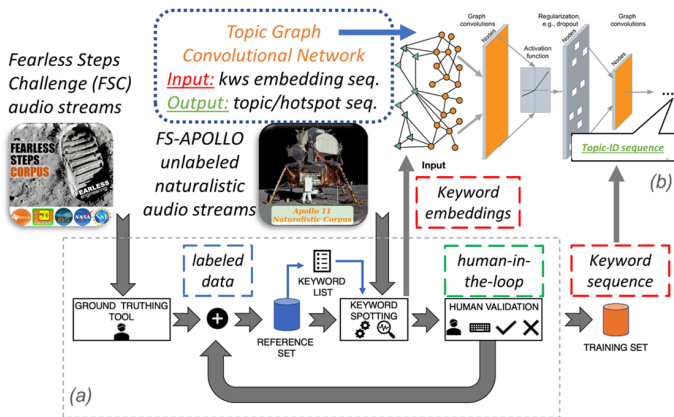
Ciertas **palabras clave** son identificativas de determinados **temas de conversación**



¡Fuerte correlación entre palabras clave y temas de conversación!

Identificación de Tema a partir de KWS

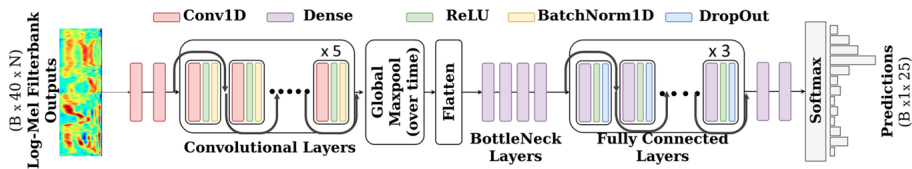
- a) Sistema de **KWS con validación humana**
- b) **Red convolucional de grafo para generar secuencias de temas** a partir de embeddings de palabras clave



A. Joglekar et al., "Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting," en 2024 NASA HRP IWS

Identificación de Tema de Extremo a Extremo

- Sí hay resultados preliminares de **identificación de tema de extremo a extremo** (es decir, ¡sin KWS!)
- Se espera obtener mejores resultados mediante la solución basada en KWS



Cuadro: Precisión (%) de identificación de tema de extremo a extremo sobre FS-P4

| Sistema | Desarrollo | Evaluación |
|------------|-------------|-------------|
| Saliency | 47,2 | 39,2 |
| ResNet-18 | 58,8 | 52,3 |
| ResNet-34 | 60,5 | 54,9 |
| S+R | 68,3 | 59,5 |

A. Joglekar *et al.*, "Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting," en 2024 NASA HRP IWS

Dr. Aditya Joglekar



- ① J. H. L. Hansen *et al.*, “Fearless Steps Apollo: Team Communications Based Community Resource Development for Science, Technology, Education, and Historical Preservation,” *en Proc. de ICASSP 2024*
- ② J. H. L. Hansen *et al.*, “Fearless Steps Apollo: Towards Community Resource Development for Science, Technology, Education, and Historical Preservation,” *en Proc. de ICASSP 2024*
- ③ A. Joglekar *et al.*, “Fearless Steps APOLLO: Identifying Conversational Mission-Critical Topics in NASA Apollo Missions Audio Based on Keyword Spotting,” *en 2024 NASA HRP IWS*
- ④ M. M. C. Shekar y J. H. L. Hansen, “Historical Audio Search and Preservation: Finding Waldo Within the Fearless Steps Apollo 11 Naturalistic Audio Corpus,” *IEEE Signal Processing Magazine*, 2023
- ⑤ A. Joglekar *et al.*, “Fearless Steps APOLLO: Challenges in keyword spotting and topic detection for naturalistic audio streams,” *The Journal of the Acoustical Society of America*, 2023
- ⑥ A. Joglekar *et al.*, “Fearless Steps: Advancements in Speech Technology and Corpus Development for Naturalistic Audio,” *en 2023 NASA HRP IWS*

Sobre el uso de IA en la curaduría de las comunicaciones de voz de las misiones Apolo de la NASA

Iván López-Espejo

I Encuentro Iberoamericano sobre Archivos e Inteligencia Artificial
Cooperación para la Preservación de la Memoria en la Era Digital

iloes@ugr.es

29 de octubre de 2024



Secretaría General
Iberoamericana
Secretaria-Geral
Ibero-Americana



RED DE ARCHIVOS
DIPLOMÁTICOS
IBEROAMERICANOS

