

**cátedra  
rtve • ugr**

**Síntesis profunda de habla  
e IA conversacional  
y sus aplicaciones**

**Presenta: Zoraida Callejas Carrión, directora de la Cátedra RTVE-UGR  
Madrid, 23 de noviembre de 2022**

**rtve**



**UNIVERSIDAD  
DE GRANADA**

**monoceros  
LABS**

# Contexto

Con el continuo crecimiento del número de servicios digitales y su complejidad, las **tecnologías del habla y la conversación** son clave para garantizar que los ciudadanos puedan **acceder de forma sencilla y personalizada** a estos servicios en su propio idioma.

Gran importancia en el ámbito periodístico: **acceder a información de actualidad y a conocimiento sobre arte, ciencia y cultura.**

# Contexto

**Escollos** para el desarrollo / adopción de tecnologías del lenguaje en España para servicios básicos a la ciudadanía:

- Necesidad de divulgar **las posibilidades de estas tecnologías**.
- Falta de **servicios a la ciudadanía innovadores** que usen estas tecnologías **en español** (dependencia de grandes empresas tecnológicas que priman el inglés y otro tipo de aplicaciones propietarias).

# Contexto

**Aportación de la cátedra** para abordar estos retos:

Investigación, innovación, divulgación y formación  
sobre tecnologías del habla e IA conversacional,  
desde una perspectiva **multidisciplinar e intersectorial** (con  
colaboración público-privada),  
buscando la **aplicabilidad** de los resultados  
a la aportación de **información fiable y fácilmente accesible** a la  
ciudadanía en español.



# La cátedra RTVE-UGR

# La Universidad de Granada



UNIVERSIDAD  
DE GRANADA

 Centros			
26 Propios	1 Adscritos	124 Departamentos	22 Investigación
 Enseñanzas ofertadas			
96 Grados	164 Másteres universitarios	28 Programas de doctorado	
 Comunidad universitaria			
55.290 Estudiantes	3.678 Personal Docente e Investigador	2.606 Personal de Administración y Servicios	

Plan  
estratégico

UGR<sup>2031</sup>

## Generación de conocimiento

que contribuya a solucionar los desafíos globales de la sociedad y que aporte valor en su entorno

Queremos ser una institución referente en investigación de excelencia, líder en un mundo cosmopolita y abierto, que impulse cambios en la sociedad a través de la generación de conocimiento, la captación del talento y el trasvase de los resultados de su actividad investigadora y creativa a la ciudadanía.

<https://canal.ugr.es/UGR2031/>



Uno de los objetivos de la UGR es la **contribución al progreso social** y al desarrollo económico de la sociedad, la **proyección de sus actividades** en el entorno más cercano y en los ámbitos nacional e internacional y el **fomento de la cooperación con el sector productivo**.

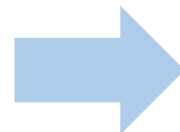
<https://mecenazgo.ugr.es/catedras-aulas/>

# Grupo de investigación en Sistemas de Diálogo Hablado y Multimodal (SISDIAL)

**20** años de experiencia en sistemas de diálogo y tecnologías del habla y el lenguaje.

Participación en numerosos proyectos de investigación y publicación en revistas y congresos de prestigio.

**Proyectos de investigación activos actualmente**



Proyecto H2020 MENHIR Mental health monitoring through interactive conversations (823907) .

Proyecto Nacional I+D+i GOMINOLA Gestión de diálogo multitarea para agentes conversacionales guiado por objetivos mediante orquestación de microservicios (PID2020-118112RB-C21).








Proyecto Nacional TED CONVERSA (comienza en diciembre).

Proyecto Andaluz I+D+i Bon-App-Petit Creación de ecosistemas basados en tecnologías móviles y lúdicas para el fomento de hábitos alimenticios y actividad física saludables entre el alumnado andaluz de educación primaria



# Equipo y datos

## Comisión mixta y comité asesor

 <p><b>ZORAIDA CALLEJAS</b> DIRECTORA DE LA CÁTEDRA</p> <p>Twitter LinkedIn GitHub</p> <p>Profesora Titular de Universidad en el Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Granada</p>	 <p><b>NIEVES ÁBALOS</b> MIEMBRO DEL COMITÉ ASESOR</p> <p>Twitter LinkedIn</p> <p>Chief Product Officer de Monoceros Labs   Conversational Interfaces Engineer   Alexa Champion &amp; AWS Community Builder</p>	 <p><b>PEDRO CÁNOVAS</b> SECRETARIO DE LA CÁTEDRA</p> <p>LinkedIn</p> <p>Jefe Ud. Coordinación Inversiones Corporación RTVE</p>	 <p><b>DAVID GRIOL</b> MIEMBRO DE LA COMISIÓN MIXTA</p> <p>LinkedIn GitHub</p> <p>Profesor Titular de Universidad en el Departamento de Lenguajes y Sistemas Informáticos de la Universidad de Granada</p>
 <p><b>CARLOS MUÑOZ</b> MIEMBRO DEL COMITÉ ASESOR</p> <p>Twitter LinkedIn</p> <p>CEO de Monoceros Labs - Innovation &amp; Digital Product Manager</p>	 <p><b>Mª CARMEN PÉREZ</b> MIEMBRO DE LA COMISIÓN MIXTA</p> <p>LinkedIn</p> <p>Subdirectora de Innovación y Estrategia Tecnológica en RTVE</p>	 <p><b>PERE VILA</b> MIEMBRO DE LA COMISIÓN MIXTA</p> <p>LinkedIn</p> <p>Director de Estrategia Tecnológica e Innovación Digital en RTVE</p>	



UNIVERSIDAD DE GRANADA



Fecha de constitución: **Junio 2022**  
Duración inicial: **2 años**

<https://catedrartve.ugr.es>

# Temáticas de la cátedra

INVESTIGACIÓN...

Síntesis  
profunda de  
habla

IA  
conversacional

Formación

Aplicada a la  
detección de  
deep fake en  
audio

Aplicada a la  
verificación y  
ampliación de  
noticias

Divulgación y  
comunicación

... APLICADA



# Síntesis profunda de habla y detección de deep fake de audio

# Síntesis profunda de habla

La síntesis profunda de habla aplica algoritmos de aprendizaje profundo para la conversación de texto a voz consiguiendo **gran naturalidad**.



- Voces inteligibles y *voice branding* para dispositivos, robots y asistentes virtuales.
- Locución de texto escrito (p.ej. lectura de webs, noticias, audio libros...).
- Audiodescripción y accesibilidad.
- Clonación de voces y voces sintéticas naturales para personas con problemas en el tracto vocal.

# Síntesis profunda de habla

**Pros:** inteligibilidad y aceptación de las voces sintéticas.

**Contras:** suplantación de la voz de una persona con fines maliciosos (**audio deep fake**).

En la **cátedra RTVE-UGR** queremos dar a conocer y poner a disposición de los periodistas herramientas que les permitan emplear la síntesis profunda en **aplicaciones innovadoras**, así como **discernir grabaciones reales de suplantaciones de identidad**.

# Deep fake de audio

- En la comunidad científica existe una larga trayectoria en la detección de suplantaciones de audio.
- **ASVspoof** (Automated speech verification spoof):
  - Énfasis en ataques a sistemas biométricos (la voz como huella digital).
  - Reto *ASVspoof challenge* para discernir de forma automática entre voces reales y sintéticas.
  - Desde 2021 el challenge tiene un apartado específico para síntesis profunda de habla.

# Deep fake de audio - Retos

- Ataque lógico (LA) vs. ataque físico (PA):
    - LA: voz sintetizada inyectada por ejemplo en una llamada.
    - PA: voz reproducida (replay attack) bien sea grabada o sintetizada.
  - Audio con “artefactos” propios de la grabación, codificación y transmisión en escenarios reales: canales telefónicos (PSTN y VoIP), reverberaciones, eco, códecs para las grabaciones...
  - Fake total y parcial / Fake de baja y alta calidad.
- No se puede generalizar la detección para cualquier escenario.
  - Necesidad de gran variedad de datos de entrenamiento y test.

# Deep fake de audio

## Cátedra RTVE-UGR

- Obtener una visión general del ámbito de estudio del deep fake en audio, tanto en la bibliografía científica como en los proyectos de investigación e innovación más relevantes.
- Obtener una visión general de las empresas del sector y su catálogo de servicios.
- Obtener una visión general de las herramientas tecnológicas disponibles para la detección de deep fake y sus características.



A person is shown from the chest up, holding a smartphone with both hands. The image is heavily overlaid with a semi-transparent red color. The person's face is partially visible, showing their eyes looking down at the phone and a slight smile. The text is centered over the phone and the lower part of the person's torso.

# IA conversacional y su aplicación en verificación de noticias

Acción: RESERVAR\_TRANSPORTE

Transporte: taxi

Hora: 18:25

Origen: Calle Mayor 40

Destino: Aeropuerto Internacional



"Llama a un taxi en 20 minutos para llevarme al aeropuerto"

Llama a un taxi en 20 minutos para llevarme al aeropuerto

Reconocimiento de habla

Procesamiento del lenguaje natural

Gestión del diálogo



De acuerdo, el taxi llegará a las 18:25 y costará 20 euros. ¿Estás conforme?

Síntesis de habla

Generación de respuesta

Acción:

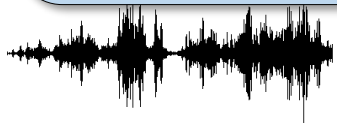
INFORMAR\_Y\_CONFIRMAR\_TRANSPORTE

Transporte: taxi

Hora: 18:25

Coste: 20€

"De acuerdo, el taxi llegará a las 18:25 y costará 20 euros. ¿Estás conforme?"



# IA conversacional aplicada al periodismo

## Asistentes al periodista

**AI4Media** - A Centre of Excellence delivering next generation AI Research and Training at the service of Media, Society and Democracy

1/9/20 - 31/8/24

<https://www.ai4media.eu/>

**Vera.ai** - Verification Assisted by Artificial Intelligence

9/2022 - 8/2025

<https://cordis.europa.eu/project/id/101070093>

# IA conversacional aplicada al periodismo

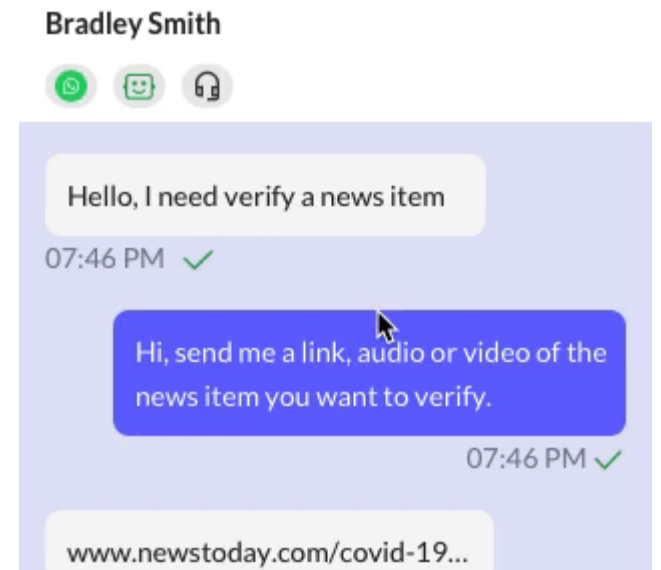
## Fact check:

- ¿Debo prestar atención a esta noticia?
- ¿Es verdadera o fake?
- ¿Es la historia completa o se han omitido partes?

Ejemplos:

<https://www.botalite.com/>

<https://www.coloq.io/en>



# IA conversacional aplicada al periodismo

## Opposing news content

- Contrarrestar las cámaras de eco en las que los usuarios sólo están expuestos a contenidos que reafirman sus opiniones previas.
- Simplemente exponer noticias con opiniones contradictorias puede reforzar opiniones aún más polarizadas. Los chatbots han obtenido resultados prometedores.

Zarouali et al. (2020) *Overcoming polarization with chatbot news? Investigating the impact of news content containing opposing views on agreement and credibility*. European journal of communication.

E.g. The New York Times' angry uncle – Agente virtual que sostiene opiniones políticas opuestas a las de los lectores, les obliga a argumentar y escuchar puntos de vista discordantes.

Hey, it's the Angry Uncle Bot. 🤖


I have LOTS of opinions.

But what kind of Uncle Bot do you want to chat with?

**SELECT YOUR REPLY**

I'm more liberal, so I'll chat with Conservative Uncle Bot.

I'm more conservative, so I'll chat with Liberal Uncle Bot.



David

Good choice. Asking a **non-threatening question** is a great way to invite Uncle Bot to discuss impeachment. People like to share their beliefs when they can do so without fear of judgment. The best questions are **open-ended, curious and non-directive**.

The impeachment? 🤖

Bah!

Just another pathetic Democratic attempt to undo the 2016 election 🤖

**HOW WOULD YOU REPLY?**

Now I remember why I unfriended you on Facebook.

O.K., I want to hear more. Where are you coming from on this?

<https://www.nytimes.com/interactive/2019/11/26/opinion/family-holiday-talk-impeachment.html>

# IA conversacional aplicada al periodismo

Tono conversacional y forma de dirigirse a la audiencia más personalizada.

- Acuerdo en la necesidad de establecer vías para la personalización de los contenidos y adquirir un tono más conversacional.
- Mejora de la accesibilidad a la información.
- Retos: mantener la línea editorial, control sobre los contenidos.

Jones and Jones. (2019) *Public service chatbots: automating conversation with BBC News*. Digital Journalism 7.

# IA conversacional aplicada al periodismo

## Periodismo participativo

- Los ciudadanos pueden aportar historias.
- Recolección de datos y multimedia.

Veglis and Kotenidis (2020) *Employing chatbots for data collection in participatory journalism and crisis situations*. Journal of Applied Journalism and Media Studies.



# IA conversacional aplicada al periodismo

## Descubrir historias

Una vez desplegados, las plataformas de desarrollo de chatbots proveen de paneles donde consultar las conversaciones mantenidas. Esto sirve:

- Desde el punto de vista de desarrollo del chatbot, para optimizar las respuestas, ver qué preguntas no obtienen respuesta, comprobar los fallos de comprensión, etc.
- Desde el punto de vista periodístico: para **descubrir nuevas historias, saber qué le interesa al público y cómo expresa ese interés con sus propias palabras.**

# IA conversacional

## Cátedra RTVE-UGR

- Conocer la potencialidad de la IA conversacional para ofrecer servicios y contenidos informativos al público.
- Conocer la potencialidad de la IA conversacional para la verificación de noticias.
- Estudiar la generación de modelos de diálogo a partir de textos informativos.



**Abiertos a colaborar**

# Colaboraciones

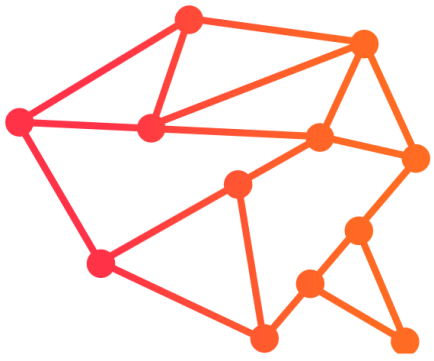
- Ecosistema de **cátedras RTVE**:  
<https://www.rtve.es/corporacion/catedras-rtve/>



- Colaboración en el contexto del **proyecto IVERES**



- **Comité asesor** de la cátedra RTVE-UGR abierto a la colaboración de otras empresas interesadas.



cátedra  
rtve • ugr

Síntesis profunda de habla  
e IA conversacional  
y sus aplicaciones  
en la verificación de noticias

Muchas gracias por su atención