

## Actividad Académicamente Dirigida (Puntuable):

# Experimentando con Herramientas y Componentes

Este trabajo consiste en investigar y presentar sobre herramientas y componentes de síntesis y reconocimiento, por lo que debe usarse software ya existente (ver abajo)

Opciones gratuitas para las componentes de TTS y ASR:

- Pocketsphinx: <https://cmusphinx.github.io/wiki/tutorialpocketsphinx/>, con modelos acústicos en español: <https://sourceforge.net/projects/cmusphinx/files/AcousticandLanguageModels/Spanish/>
- Festival: <http://www.cstr.ed.ac.uk/projects/festival/>, con una voz en español: <https://sangonz.wordpress.com/2010/05/22/spanish-voices-for-festival/>
- API de IBM Watson, de uso gratuito para pruebas.
- Hay varias opciones más de servicios en la nube. Son libres de usar las que prefieran.

### Modalidad de entrega

- El trabajo se debe realizar en grupos de dos a tres integrantes.
- **Antes del 1 de diciembre** (y cuanto antes, mejor): Enviar una breve descripción del dominio elegido a [alfredo.marquez@dti.uhu.es](mailto:alfredo.marquez@dti.uhu.es), para ver si está bien y poder continuar con la presentación y la implementación.
- **A partir del viernes 15 de diciembre** los autores deberán realizar una presentación oral y demostración del trabajo en clase, en su ordenador personal. Esta presentación durará entre 10 y 15 minutos, durante los cuales se deberá explicar el dominio, mostrar las funcionalidades del sistema y describir su implementación. La fecha límite de exposición será los días de clase teórica.
- Además, se deberá entregar un breve informe, en formato PDF y de cuatro (4) carillas como máximo, que describa los mismos aspectos de la presentación oral. Este informe debe enviarse a [alfredo.marquez@dti.uhu.es](mailto:alfredo.marquez@dti.uhu.es). La fecha límite de entrega del informe es el día del examen. Las Actividades Dirigidas entregados fuera de término serán aceptados, pero la demora incidirá negativamente en la nota.

### Modo de evaluación

- (1) El trabajo se evaluará desde la óptica de la exposición clara de los conceptos por parte de los oradores.

- (2) En el caso de implementar alguna demo con componentes en una pequeña aplicación también se valorará. No se evaluarán aspectos estéticos de la interfaz, en este caso sino estrictamente los relativos al uso del habla para interactuar con el sistema.
- El TP tiene una nota máxima de 10. Cumplir con los objetivos descritos en el punto (1) otorga 7 puntos. Los restantes 3 puntos corresponden al punto (2) en cuanto a la implementación de la interacción mediante el habla, según la opinión del profesor