



# MediaEval



YESID OSPITIA MEDINA  
ASESORÍA TECNOLÓGICA INTEGRAL

# Agenda

- 1. DEAM Dataset**
2. Estructura general
3. Detalle de las carpetas
4. Instrumento de etiquetado emocional



# 1. DEAM Dataset

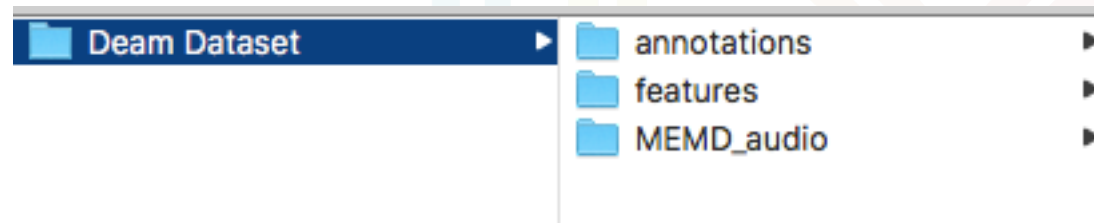
- “DEAM dataset consists of 1802 excerpts and full songs annotated with valence and arousal values both continuously (per-second) and over the whole song”.
- Es una de las bases de datos más ampliamente utilizadas en la disciplina del reconocimiento de emociones en la música.
- Es de acceso abierto, y una gran contribución a la comunidad científica.

# Agenda

1. DEAM Dataset
- 2. Estructura general**
3. Detalle de las carpetas
4. Instrumento de etiquetado emocional

# Estructura general

Carpeta	Contenido	Tipo de archivos
<b>annotations</b>	Anotaciones de tipo arousal-valence (AV) sobre cada una de las canciones.	Extensión .csv
<b>features</b>	Características (features) extraídas de cada archivo de sonido	Extensión .csv
<b>MEMD_audio</b>	1802 archivos de sonido.	Extensión .mp3



URL: <http://cvml.unige.ch/databases/DEAM/>

# Carpeta features

La carpeta presenta en su contenido 1.802 archivos en formato csv, cada uno de estos archivos corresponde a las features extraídas de uno de los archivos contenidos en la carpeta MEDMD\_Audio (archivos de audio).

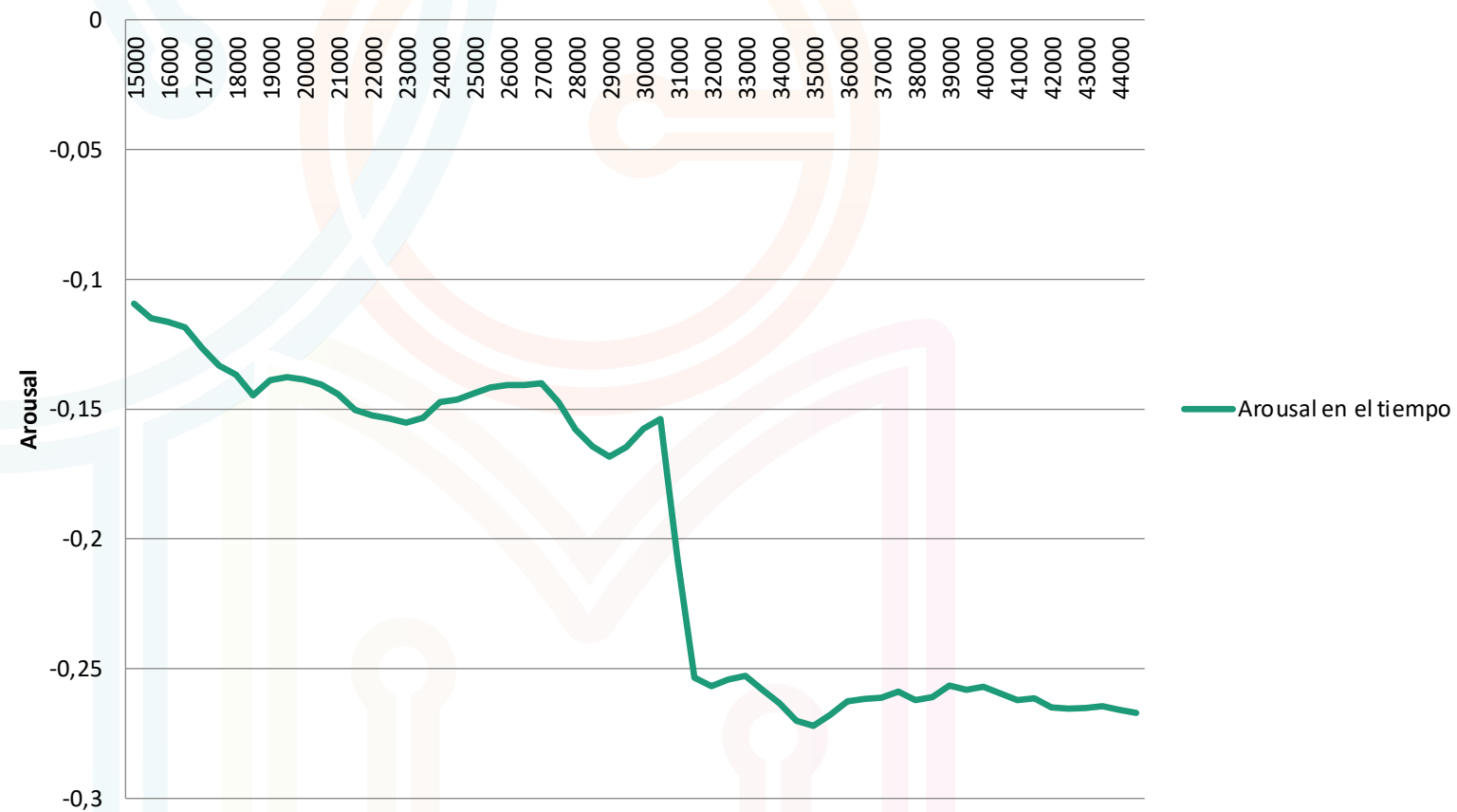
A continuación algunas características relevantes del archivo csv de features:

- Cuenta con 449 muestras de features, dichas muestras son tomadas en espacios de 0,5 ventanas de tiempo (frametime). En cada momento de muestra (cada frametime) se toman los valores de todos los features analizados en la fila 1 del archivo.
- El archivo maneja un total de 260 features.

# Carpeta annotations

- **annotations averaged per song:** Incluye la información relativa a los valores medios calculados de acuerdo a todos las anotaciones individuales de los evaluadores (**annotations per each rater**) Los archivos que encontramos en esta carpeta corresponden a:
  - **dynamic (per second annotations)/arousal.csv:** En este archivo se puede evidenciar la identificación de un valor de arousal para una canción determinada. Dicha medición de arousal se hace por primera vez a los 1500 ms de la canción y luego a partir de allí la medición se vuelve a realizar cada 500 ms, hasta llegar a la última medición que se presenta en 44.500 ms. Los primeros 15 segundos se omiten por cuestiones de estabilización en el proceso de anotación durante el inicio de la canción. Para cada archivo de sonido, existe una fila en este archivo, en donde se pueden evidenciar las mediciones de arousal. En la siguiente figura se presenta la variación del arousal para la canción 2.mp3

## Arousal "2.pm3"

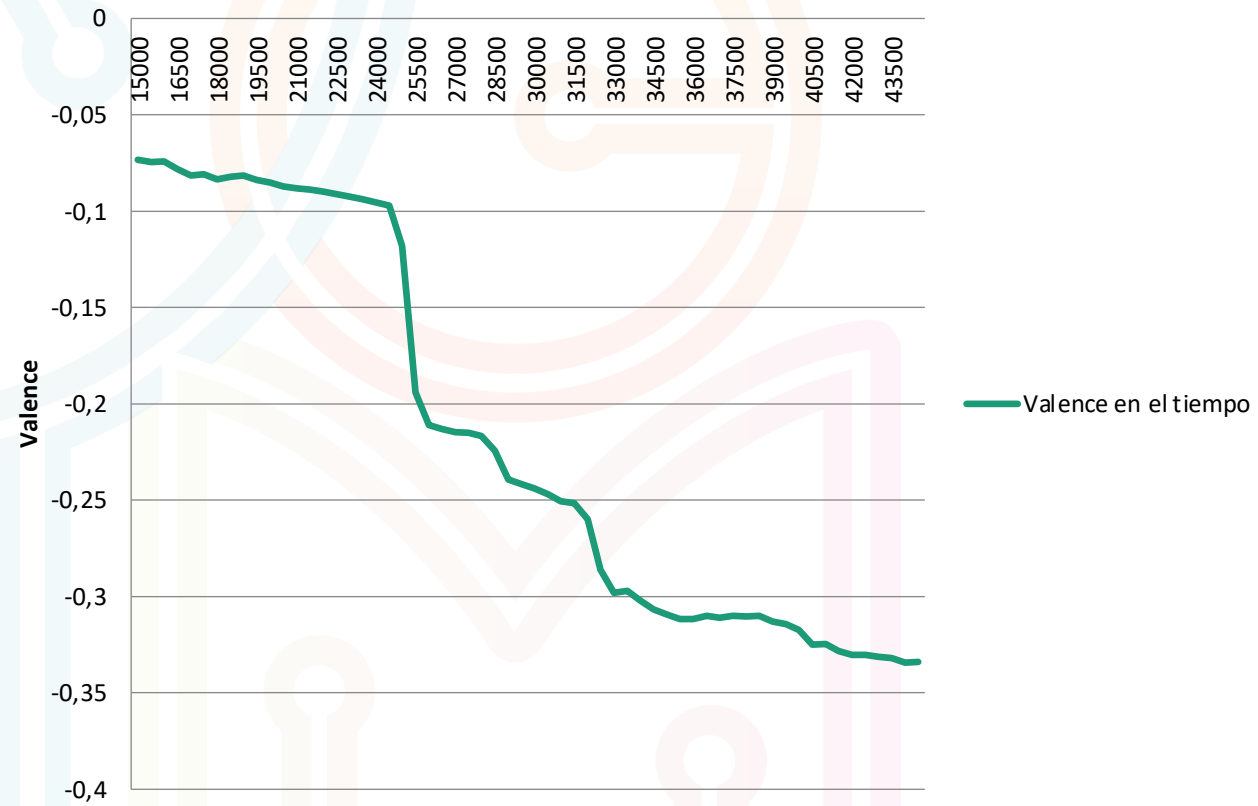




# Carpeta annotations

**dynamic (per second annotations)/valence.csv:** En este archivo se puede evidenciar la identificación de un valor de valence para una canción determinada. Dicha medición de valence se hace por primera vez a los 1500 ms de la canción y luego a partir de allí la medición se vuelve a realizar cada 500 ms, hasta llegar a la última medición que se presenta en 44.500 ms. Para cada archivo de sonido existe una fila en este archivo, en donde se pueden evidenciar las mediciones de valence. En la siguiente figura se presenta la variación del valence para la canción 2.mp3

## Valence "2.pm3"



# Carpeta annotations

- **song\_level/static\_annotations\_averaged\_songs\_1\_2000.csv:** Este archivo solo aplica para las primeras 1744 canciones (archivos mp3) contenidos en la carpeta MEDMD\_Audio. Este archivo presenta los valores medios para el valence y arousal, como también respectivamente sus valores de desviación. Para el caso del archivo 2.mp3 tenemos un valence medio 3,1 con std 0,94 y un arousal medio 3,0 con std 0,63.
- **song\_level/ static\_annotations\_averaged\_songs\_2000\_2058:** Este archivo solo aplica para las últimas 58 canciones (archivos mp3) contenidos en la carpeta MEDMD\_Audio. En este archivo al igual que en **static\_annotations\_averaged\_songs\_1\_2000.csv** se presenta el valor medio para valence y arousal, como también las desviaciones respectivas. Por otra parte, la información de este archivo es más completa al presentar también los valores máximos y mínimos tanto para la media y std del valence y el arousal .



# Carpeta annotations

**annotations per each rater:** En los archivos presentados en esta carpeta se encuentra las anotaciones de cada uno de los evaluadores sobre cada una de las canciones. El proceso de anotación es dinámico, lo que se conoce como *over-time*. **En este caso, no tenemos vectores promediados.**

# Instrumento para etiquetado emocional

The interface consists of several components:

- Music Player:** Located at the top left, it shows a play button, a progress bar at 00:06 of 00:10, and a volume icon.
- Slider:** Below the player, a horizontal scale from -10 to 10 with a red oval marker positioned at 7. A minus sign is at -10 and a plus sign is at 10.
- Player Instructions:** A pink box on the right with a close button (X) and the following text:
  - Play button is disabled
  - Press **p** to Play/Pause, **Ctrl+q** to Replay current music
  - Audio will not play unless mouse is in the box
  - Moving mouse outside box will pause annotation
- Show instructions:** A blue button below the instructions box.
- Annotating Arousal progress:** A green progress bar below the button, labeled "Annotating Arousal progress: 1 / 3".
- Arousal 7.07:** Text displayed below the slider.
- What do you think the overall arousal of this song is:** A question above a scale of nine self-assessment manikins, numbered 1 to 9. The manikins show increasing levels of arousal from a neutral face to a highly expressive, energetic face.
- How confident are you about the annotation you just provided?** A question above a set of radio buttons with the following options: Not at all, Not really, Don't know, Somewhat, Very much.
- Disclaimer:** A line of text stating: "Your response to this question will not be used as a basis for acceptance or rejection of your HIT and will be solely used for our statistical analysis."

Fig 2. Annotation interface for both continuous (upper-left corner) and static per song (middle; using the self-assessment manikins [43]) ratings of arousal.



Gracias!



YESID OSPITIA MEDINA  
ASESORÍA TECNOLÓGICA INTEGRAL