

# Reporte del desarrollo de una Ontología

Saul Ivan Rivas Vega  
Inteligencia Artificial

25 de septiembre de 2019

## 1. Introducción

Una definición para ontología es una descripción formal y explícita de conceptos, sus propiedades y las relaciones entre ellos. [1]. En este trabajo se presenta 'Audio Subset'. Una ontología para la clasificación de eventos acústicos.

Actualmente ha proliferado el interés de realizar un análisis en los distintos eventos sonoros en el ambiente con varios propósitos tales como el determinar el contexto para sistemas de agentes o una simple identificación de eventos en un extracto de sonido [2]. Así mismo existen una serie de tareas en las que se utilizan métodos de clasificación con uso de deep learning [3]. Sin embargo es frecuente encontrarse con distintas categorizaciones para dichos eventos, casi una distinta con cada nuevo trabajo dependiendo el dominio [4, 5, 6].

Esto llevo a proponer una ontología al no existir una estándar. El trabajo resultante fue 'Audio Set' [7]. De la cuál se obtuvo un subconjunto de las categorías para desarrollar 'Audio Subset', se describirá las diferencias con respecto a 'Audio Set' y algunos conceptos adicionales.

## 2. Desarrollo

Buscaremos definir los siguientes conceptos [8]:

**Clase** Conjunto de elementos con caracteres comunes.

**Propiedad** Atributo o cualidad esencial de alguien o algo.

**Característica** Que da carácter o que sirve para distinguir a alguien o algo de sus semejantes.

Ahora revisemos el subconjunto de 2 niveles de 'Audio Set':



Figura 1: Subconjunto como se muestra en [7].

Ademas de definir el 3er nivel y posteriores, se realizaron las siguientes modificaciones a los 2 primeros niveles:

- Cambiar el nombre de 'Hands' por 'Limbs', para generalizar el hecho de que el sonido puede provenir de la interacción de las extremidades y no solo de las manos.
- Dejar fuera 'Whistling' pues este puede entrar como instancia en alguna de las otras clases.
- Dejar fuera 'Octoacoustic emissions' pues realmente no es sonido detectable de manera externa (solo de manera interna en el oído).
- Dejar fuera 'Music Concepts' pues su definición es mas subjetiva ya que en la mayoría de la literatura se trabajan como características del sonido.
- Dejar fuera 'Music Mood' puesto que es determinado por una evaluación subjetivo de la emoción evocada por una pieza musical.
- Agregar 'Steel' a 'Sound of things' debido a que puede ser categorizado de igual forma que los demás materiales en la clase.
- Dejar fuera 'Miscellaneous sources' pues las entidades categorizadas carecían de similitudes significantes para ser agrupadas.
- Dejar fuera 'Specific Impact' pues las subclases no se diferenciaban mucho de 'Generic impact sounds' de la clase 'Source-ambiguous sounds'.
- Dejar fuera 'Onomatopoeia' pues son sonidos que hacen referencia a otros.
- Dejar fuera 'Silence' y 'Other sourceless' puesto que no se refieren a un sonido detectable.

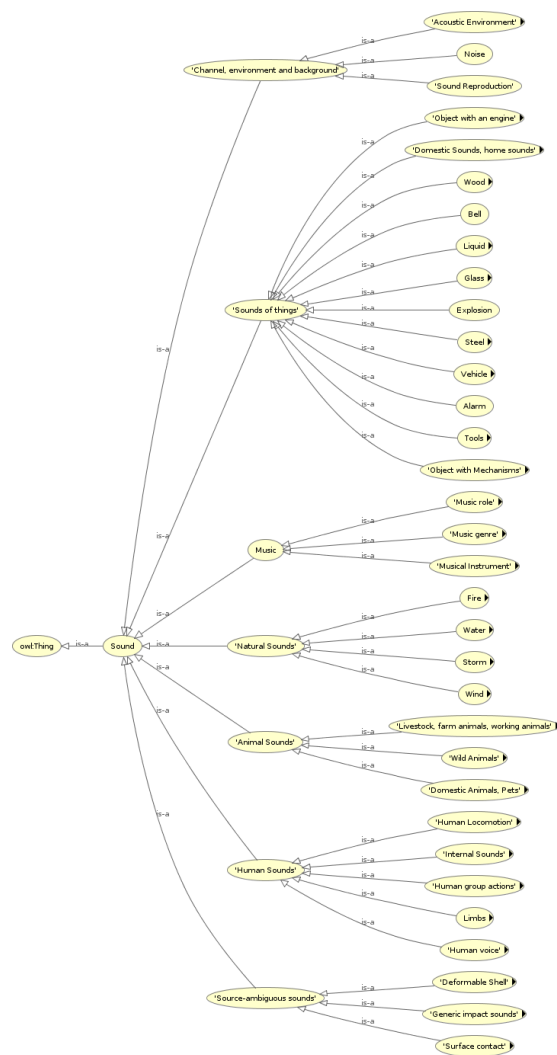


Figura 2: Subconjunto de 2 niveles de Audio Subset.

### 3. Conclusiones

El objetivo en el desarrollo de 'Audio Subset' como una exploración de las distintas categorizaciones en eventos acústicos se cumplió además de detectar las distintas modificaciones con respecto a la tarea a realizar. Se dejaron fuera varios eventos acústicos que no necesariamente son detectables, se entiende que 'Audio Set' los incluye pues es de una categorización mas general, y 'Audio Subset' trata de especificar las clases que sirvan para una aplicación práctica. Finalmente el desarrollo de una ontología tiene gran valor como ejercicio en el aprendizaje sobre la representación del conocimiento con la finalidad de llevar a cabo una investigación en el dominio representado.

### Referencias

- [1] N. F. Noy and D. L. McGuinness, "Ontology Development 101: A Guide to Creating Your First Ontology," p. 25.
- [2] T. Virtanen, M. Plumbley, and D. Ellis, *Computational analysis of sound scenes and events*, 1st ed. New York, NY: Springer Science+Business Media, 2017.
- [3] H. Purwins, B. Li, T. Virtanen, J. Schlüter, S. Chang, and T. Sainath, "Deep Learning for Audio Signal Processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 2, pp. 206–219, May 2019. [Online]. Available: <https://arxiv.org/pdf/1905.00078.pdf>
- [4] D. Eck and J. Schmidhuber, "Finding temporal structure in music: blues improvisation with LSTM recurrent networks," in *Proceedings of the 12th IEEE Workshop on Neural Networks for Signal Processing*. Martigny, Switzerland: IEEE, 2002, pp. 747–756. [Online]. Available: <http://ieeexplore.ieee.org/document/1030094/>
- [5] N. Turpault, R. Serizel, and E. Vincent, "Semi-supervised Triplet Loss Based Learning of Ambient Audio Embeddings," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton, United Kingdom: IEEE, May 2019,

- pp. 760–764. [Online]. Available: <https://ieeexplore.ieee.org/document/8683774/>
- [6] E. Benetos, G. Lafay, M. Lagrange, and M. D. Plumbley, “Detection of overlapping acoustic events using a temporally-constrained probabilistic model,” in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Shanghai: IEEE, Mar. 2016, pp. 6450–6454. [Online]. Available: <http://ieeexplore.ieee.org/document/7472919/>
  - [7] J. F. Gemmeke, D. P. W. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, “Audio Set: An ontology and human-labeled dataset for audio events,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New Orleans, LA: IEEE, Mar. 2017, pp. 776–780. [Online]. Available: <https://ai.google/research/pubs/pub45857.pdf>
  - [8] R. A. Española and A. de Academias de la Lengua Española, *Diccionario de la lengua española*, vigesimotercera edición, edición del tricentenario ed. Real Academia Española, 2014. [Online]. Available: <http://dle.rae.es/?id=DgIqVCc>