



## Becas colaboración curso 2023/2024

Fecha: 29 Mayo 2023

### Vicerrectorado de Investigación

Subcomisión de I+D+i

Propuesta del departamento *LINGÜÍSTICA APLICADA*

**Núm Proyecto: 2023/13/00001**

#### Responsable

Carrió Pastor, María Luisa

#### E-mail

lcarrio@idm.upv.es

#### Ext.

75381

#### Responsable

Periñán Pascual, José Carlos

#### E-mail

jopepas3@upvnet.upv.es

#### Ext

49433

#### Título proyecto

ANÁLISIS DE TEXTOS CIENTÍFICOS MEDIANTE GRANDES MODELOS LINGÜÍSTICOS (LLM)

#### Valoración proyecto

4

#### Descripción proyecto

El objetivo del proyecto a desarrollar es el estudio de textos desde el punto de vista de los grandes modelos lingüísticos. Los grandes modelos lingüísticos (LLM) es un modelo de lenguaje que consiste en una red neuronal con muchos parámetros (generalmente miles de millones de palabras o más), entrenados en grandes cantidades de texto sin etiquetar mediante aprendizaje autorevisado o aprendizaje semirevisado. Estos análisis suponen un paso más que la mera descripción estadística de los elementos del discurso narrativo, sino que se considera la estructura sistémica del mismo desde el punto de vista narrativo. En particular se estudiará la estructura sistémica de artículos científicos en diversas disciplinas, así como otro tipo de corpus. También se introducirán medidas del estudio de la complejidad de los mismos, tales como la entropía. Todo ello permitirá conocer de forma más exacta la composición y patrones de las lenguas, específicamente, del español y del inglés.

#### Actividades a realizar por el alumno

El alumno aplicará la metodología con el fin de estudiar la estructura de textos científicos desde el punto de vista de los grandes modelos lingüísticos, analizando la estructura gramatical como la carga semántica del discurso. Tras un preprocesado sencillo, los alumnos trabajarán con are Common Crawl, The Pile, MassiveText, etc. Para el desarrollo del proyecto en la parte de grandes modelos lingüísticos, se cuenta con la colaboración del Prof. J. Alberto Conejero del Instituto Universitario de Matemática Pura y Aplicada (IUMPA-UPV).

#### Localización de la actividad (Campus)



## Becas colaboración curso 2023/2024

*Fecha: 29 Mayo 2023*

Vera

### **Horario**

Tardes o mañanas