



Becas colaboración curso 2022/2023

Fecha: 01 Junio 2022

Vicerrectorado de Investigación

Subcomisión de I+D+i

Propuesta del departamento *SISTEMAS INFORMÁTICOS Y COMPUTACIÓN*

Núm Proyecto: 2022/32/00016

Responsable

Hurtado Oliver, Lluís Felip

E-mail

lhurtado@dsic.upv.es

Ext.

83500

Responsable

Sanchís Arnal, Emilio

E-mail

esanchis@dsic.upv.es

Ext

73510

Título proyecto

Aplicación de Técnicas de Procesamiento del Lenguaje Natural para la Recuperación de Información basada en representaciones vectoriales densas

Valoración proyecto

4

Descripción proyecto

El grupo de investigación ELiRF tiene una dilatada experiencia en distintas aplicaciones del Procesamiento de Lenguaje Natural. Algunas de estas aplicaciones son la monitorización de redes sociales, el análisis de artículos de prensa y programas de debate televisivos, y el resumen automático. Actualmente está trabajando dentro de un proyecto del Ministerio de Ciencia e Innovación (ref: PDC2021-120846-C44, "DESARROLLO DE UN PROTOTIPO PRECOMPETITIVO PARA EL ANALISIS AFECTIVO DE INFORMACION MULTIMEDIA"), consistente en el desarrollo de un sistema de Recuperación de Información sobre una colección de documentos audio-visuales.

El grupo de investigación dispone de gran cantidad de recursos hardware, software y corpus que darán soporte durante la realización del trabajo asociado a la beca.

Actividades a realizar por el alumno

El candidato o candidata se integrará en el grupo de investigación, que forma parte del instituto de investigación VRAIN, asistiendo a reuniones y colaborando con sus miembros en los distintos proyectos activos.

En particular trabajará en la construcción de un sistema de recuperación de información sobre una colección de documentos audiovisuales para los que se dispone de la transcripción manual y automática. Para la construcción del índice de la colección tanto los documentos como las consultas se representarán mediante vectores densos. Se abordarán diferentes aproximaciones para la obtención de la representación de las consultas y documentos en base a embeddings incontextuales y contextuales y para el cálculo de su similitud. Con todo ello, el becario o becaria aprenderá los fundamentos teóricos (modelos de lenguaje, métricas de evaluación, etc.) y tecnológicos (word embeddings, deep learning, etc.) que serán necesarios para abordar las tareas encomendadas.



Becas colaboración curso 2022/2023

Fecha: 01 Junio 2022

Localización de la actividad (Campus)

Campus de Vera

Horario

3 horas diarias, a acordar con el o la estudiante