

La inteligencia artificial (IA) aplicada a la localización, investigación y gestión de fondos manuscritos digitalizados de archivos históricos: el proyecto Carabela

24/10/19



Instituto Andaluz del Patrimonio Histórico
CONSEJERÍA DE CULTURA Y PATRIMONIO HISTÓRICO

La aplicación de las técnicas de Inteligencia Artificial (IA) está abriendo nuevas perspectivas de trabajo y desarrollo en diferentes ámbitos científicos. Desde noviembre de 2018 el Instituto Andaluz del Patrimonio Histórico (IAPH), a través de su Centro de Arqueología Subacuática (CAS), está trabajando en el desarrollo del *Proyecto Carabela: Indexación probabilística de colecciones de manuscritos para protección del patrimonio histórico subacuático*, una iniciativa liderada por el Centro de Investigación de Reconocimiento de Patrones y Tecnología del Lenguaje Humano (PRHLT) de la Universidad Politécnica de Valencia, desarrollada en base a la ayuda concedida por la Fundación BBVA en la convocatoria de 2018 para equipos de investigación en Humanidades Digitales.

Si bien las técnicas de Reconocimiento Óptico de Caracteres (OCR) han evolucionado hasta permitir la búsqueda automatizada a texto completo en documentos impresos, estas no son de utilidad cuando se aplican a documentos manuscritos de diferentes grafías. La finalidad del proyecto Carabela es analizar el potencial

de la IA desde las técnicas de reconocimiento formal como estrategia para la localización, investigación y gestión de documentos manuscritos en fondos digitalizados de archivos históricos.

Para el desarrollo del proyecto se ha seleccionado, a modo experimental, una colección documental compuesto por 150.000 imágenes (de interés para la arqueología subacuática por el objeto del proyecto) pertenecientes a los fondos del Archivo General de Indias y el Archivo Histórico Provincial de Cádiz. Sobre ella se ha aplicado una nueva metodología de aprendizaje automático que permite indexar probabilísticamente palabras en texto manuscrito, permitiendo, además, realizar búsquedas temáticas en colecciones masivas de documentos históricos desde diferentes intereses de investigación y/o gestión. De esta manera se ha podido, además, localizar e investigar de forma eficaz información sobre los naufragios históricos, así como clasificar los expedientes en base a las diferentes temáticas de su contenido.

Organiza

Instituto Andaluz de Patrimonio Histórico (IAPH)

Colabora

Asociación de Archiveros de Andalucía

Sede

IAPH, Sala de Conferencias, Edificio A. Camino de los Descubrimientos s/n. Isla de la Cartuja, Sevilla

Fecha

jueves 24 de octubre de 2019

Horario

de 16:30 – 18:30 h

Destinatarios

Personal investigador y profesionales de los Archivos

Inscripción

Será necesario el envío de un correo electrónico a la dirección phforma.iaph@juntadeandalucia.es con los datos personales y profesionales informando de su interés por la asistencia a la jornada.

Admisión por orden de llegada, hasta completar aforo.

Actividad de carácter gratuito.



Programa

16:30-16:50 h.

El proyecto Carabela: origen, objetivos y demostración de uso.
Carlos Alonso Villalobos (IAPH).

16:50-17:30 h.

Las tecnologías de la inteligencia artificial (IA) aplicadas a textos manuscritos: evolución, desarrollo y proyectos emblemáticos.
Enrique Vidal Ruiz (PRHLT).

17:30-18:00 h.

Desarrollo del proyecto.
Preparación, organización y tratamiento de la documentación.
Alejandro H. Toselli y Verónica Romero (PRHLT).

18:00-18:30 h.

Mesa redonda.



En colaboración con:



Instituto Andaluz del Patrimonio Histórico
CONSEJERÍA DE CULTURA Y PATRIMONIO HISTÓRICO

**INSTITUTO ANDALUZ
DEL PATRIMONIO HISTÓRICO**

Camino de los Descubrimientos, s/n.
Isla de la Cartuja. 41092 Sevilla
Tel.: + 34 955 037 047 Fax: + 34 955 037 001
cursos.iaph@juntadeandalucia.es
www.iaph.es

