

NÚMERO DE HORAS: 22,5 h.

LUGAR DE CELEBRACIÓN:

UNIVERSIDAD DE BURGOS
BIBLIOTECA UNIVERSITARIA
Aula de Formación
Plaza Infanta Doña Elena s/n
09001 Burgos

IMPORTE DE LA MATRÍCULA:

Los colectivos que pueden acogerse a la modalidad de matrícula reducida se pueden consultar en el apartado [tipos de matrícula](#)

- Ordinaria: 100 euros
- Reducida: 60 euros
- Gratuita

NOTAS DE INTERÉS:

- Podrán acogerse a la matrícula reducida los miembros y colaboradores del equipo PROTEO.

INFORMACIÓN GENERAL:

- Para el estudiantado matriculado en los Títulos Oficiales adaptados al Espacio Europeo de Educación Superior (Grados de la UBU) se reconocerá **0,5 créditos por cada curso o actividad equivalentes a 12,5 h.**

Información:

Servicio de Estudiantes y Extensión Universitaria
Edificio de Administración y Servicios
C/ Don Juan de Austria, 1 – 09001 Burgos
Tfnos: 947 25 80 80 / 947 25 87 00
correo: cverano@ubu.es
<http://www.ubu.es/cverano>

Curso patrocinado por:



unir LA UNIVERSIDAD
EN INTERNET



Organiza



**UNIVERSIDAD
DE BURGOS**

Patrocina



Excma. Diputación
Provincial de Burgos



**FILOLOGÍA DIGITAL E
INTELIGENCIA ARTIFICIAL:
TRANSCRIPCIÓN
AUTOMÁTICA,
ESTILOMETRÍA
Y ESCANSIÓN
COMPUTARIZADA**

Dirección:
Dra. María Luisa Lobato
Codirección:
Dr. Álvaro Cuéllar

**del 4 al 6 de septiembre
BURGOS**



**UNIVERSIDAD
DE BURGOS**

OBJETIVOS

El curso ofrece a los asistentes la posibilidad de conocer de forma teórica y práctica las investigaciones y procesos digitales y de Inteligencia Artificial que se están desarrollando recientemente en el ámbito de la Filología. Entre estos destaca la aplicación de IA para la transcripción automática de documentos antiguos, la Estilometría para relacionar las obras literarias por sus usos léxicos y el uso de Python para el manejo informático de los textos del interés del investigador, así como del análisis rítmico de las obras poéticas. Se desglosan aquí los objetivos particulares del curso:

- Presentar las dificultades para la transcripción automática de documentos antiguos.
- Mostrar cómo transcribir documentos antiguos (impresos y manuscritos) a través de Inteligencia Artificial.
- Presentar de qué manera construir un detector de copistas o imprentas con Inteligencia Artificial.
- Presentar el problema filológico de la atribución textual de textos literarios.
- Mostrar la importancia de la modelización y manejo de un corpus.
- Aproximación al *data mining*: extracción de textos de la web.
- Aproximación al tratamiento de datos: del html al txt vía regex (*Regular expressions*).
- Introducción al XML/TEI. Codificación de una obra de teatro/poesía.
- Trabajar con transformación / extracción del XML al TXT.
- Explicar el almacenamiento, recuperación y gestión de datos. Github (control de cambios)
- Introducir a los métodos de análisis estilísticos con ayuda de ordenadores.
- Introducción teórico-práctica a la estilometría.
- Conocer lenguajes de programación. Vocabulario básico. Lenguaje R. Funciones, variables, paquetes, ayuda.
- Presentar y profundizar en Stylo, Metrics: Cluster, PCA, Bootstrap, etc.
- Métodos de análisis secuencial (estudio de caso de comedias colaboradas).
- Introducir a humanistas, en especial a filólogos, en los métodos que están detrás del análisis estadístico: PCA, *Machine Learning*, *Neural Networks*, AI.
- Presentar el análisis y visualización con redes (Gephi).
- Dar a conocer Python como herramienta para el estudio de los textos.

- Mostrar las problemáticas en torno al estudio rítmico de la poesía.
- Enseñar cómo pueden construirse algoritmos para escandir automáticamente textos poéticos.
- Aproximarse a técnicas de Inteligencia Artificial y *machine learning* para la clasificación autorial.

DIRIGIDO A:

Titulados y estudiantes interesados en Humanidades Digitales. Tendrán prioridad los filólogos y los informáticos.

Lunes, 4 de septiembre

09:30 a 10:30 h. Introducción

10:30 a 11:30 h. La investigación sobre teatro del Siglo de Oro: a grandes males, remedios digitales
Dr. GERMÁN VEGA GARCÍA-LUENGOS

11:30 a 12:00 h. Instalación de programas (opcional)

12:00 a 12:30 h. Pausa café

12:30 a 14:00 h. Introducción a Transkribus como sistema de IA para la transcripción automática
Dr. ÁLVARO CUÉLLAR

14:00 a 16:00 h. Comida

16:00 a 17:30 h. Usos avanzados de Transkribus: modernización ortográfica, detección estructural
Dr. ÁLVARO CUÉLLAR

17:30 a 18:00 h. Pausa café

18:00 a 19:30 h. Introducción a la estilometría
Dra. LAURA HERNÁNDEZ

Martes, 5 de septiembre

09:30 a 10:30 h. Edición crítica digital y anotación de corpus para estudios literarios computacionales
Dr. BORJA NAVARRO

10:30 a 12:00 h. Aplicación de métodos estilométricos con stylo
Dra. LAURA HERNÁNDEZ

12:00 a 12:30 h. Pausa café

12:30 a 14:00 h. Análisis secuencial con stylo
Dra. LAURA HERNÁNDEZ

14:00 a 16:00 h. Comida

16:00 a 17:30 h. Análisis de redes estilométricas con stylo y Gephi
Dra. LAURA HERNÁNDEZ

17:30 a 18:00 h. Pausa café

18:00 a 19:30 h. Actividad extra

Miércoles, 6 de septiembre

09:30 a 10:30 h. Introducción a XML/TEI. Codificación de una obra de teatro/poesía
Dr. SIMON KROLL

10:30 a 12:00 h. Introducción a Python como entorno de trabajo
Dr. SIMON KROLL

12:00 a 12:30 h. Pausa café

12:30 a 14:00 h. Presentación del estado de la cuestión sobre los usos rítmicos en el Siglo de Oro
Dr. SIMON KROLL

14:00 a 16:00 h. Comida

16:00 a 16:30 h. Proyectos

16:30 a 17:30 h. Uso de machine learning para la clasificación autorial de obras literarias en función de sus ritmos
Dr. SIMON KROLL

17:30 a 18:00 h. Cierre