



## Memoria de propuesta de Trabajo de Fin de Título de la Escuela de Ingeniería Informática de la ULPGC

**Descripción, objetivos y metodología del trabajo:** Se describirá el contexto donde se va a desarrollar el trabajo y la situación, problema o aplicación concreta que se pretende realizar, así como los objetivos que se pretenden alcanzar y la metodología a seguir. En el caso de Doble Grado, especificar con claridad la relación con los módulos y las materias de ambas titulaciones.

### 1. Descripción (obligatorio): Descripción/motivación del trabajo

Dada la nueva ley de protección de datos, la publicación y el uso de datos referentes a personas en documentos tanto oficiales como públicos puede violar sus derechos de privacidad. Por ello hace falta alguna herramienta que haga posible la anonimización/ofuscación/disgregación de datos personales de acuerdo con esta nueva ley.

### 2. Objetivos (obligatorio): Objetivos/resultados del trabajo

El objetivo que persigue este proyecto será el de eliminar o reducir al mínimo los riesgos de identificación de los titulares de los datos. Se trata de construir una herramienta con la que se podrá reducir los riesgos de identificación asociados a datos personales en diferentes tipos de documentos. Es decir, se pretende crear un sistema cuya entrada sea documentos (docx, pdf, txt, hojas de cálculos...) y la salida será la ofuscación/disgregación/anonimización de datos personales, en este caso nombres propios, apellidos y DNI.

### 3. Metodología (obligatorio): Cómo se va a desarrollar, herramientas a usar, tecnologías

Se desarrollará aplicando técnicas de procesamiento de lenguaje natural, y se hará uso de librerías de Python relacionadas con este ámbito. De la misma forma se utilizarán herramientas aplicadas a la creación de servicios web.

Durante el desarrollo se contará con el asesoramiento de una empresa dedicada a la protección de datos, con lo cual se llevarán a cabo diversas entrevistas a fin de validar todos los objetivos de este proyecto.

Medios específicos (en su caso):

Justificación de la necesidad de varios tutores (en su caso):

**Competencias desarrolladas:** Se asume que todos los TFT incluyen las competencias generales recogidas en el Guía Docente de la asignatura correspondiente. Deberá elegirse a continuación **al menos una competencia específica** (una por titulación en el caso del doble grado) que esté más directamente relacionada con el trabajo. Se pueden consultar en los siguientes enlaces según corresponda: [GII](#) (pág.6-10), [MII](#) (pág.2-3) y [GADE](#).

Código/s: CP01, CP03, CP04

Firmas (en caso de que no se haya firmado digitalmente):

**Plan de trabajo:** Se desglosará de manera detallada el trabajo del TFT en fases, con su duración estimada en horas (total 300). Cada fase, a su vez, se desglosará en tareas concretas; en el caso del doble grado deberá indicarse para cada tarea si es común o específica de alguno de los grados.

<b>Fases</b>	<b>Duración Estimada (horas)</b>	<b>Tareas (nombre y descripción, obligatorio al menos una por fase)</b>
Estudio previo / Análisis	25	Tarea 1.1: Estudio del estado del arte.
		Tarea 1.2: Estudio sobre las técnicas de procesamiento natural.
		Tarea 1.3: Estudio sobre el procesamiento de lenguaje natural a la búsqueda de entidades (NER) en un texto.
Diseño / Desarrollo / Implementación	215	Tarea 2.1: Realización de una herramienta mediante Python que permita la lectura/escritura de textos en diferentes formatos como docx, pdf, txt, hojas de cálculo...
		Tarea 2.2: Desarrollo del código que permita la detección de entidades
		Tarea 2.3: Integrar las aplicaciones/herramientas anteriores en un servicio web.
Evaluación / Validación / Prueba	20	Tarea 3.1: Estudiar el rendimiento de las técnicas de procesamiento de lenguaje natural utilizadas.
		Tarea 3.2: Estudiar el rendimiento del servicio.
Documentación / Presentación	40	Tarea 4.1: Realizar la memoria correspondiente de todo el proceso.

(añadir más filas si es necesario)

NOTA: la estructura aquí reflejada es un simple ejemplo. Deberá adaptarse a las peculiaridades de cada trabajo y ser coherente con lo que se haya indicado en la metodología.

- El proponente de este TFT manifiesta que el trabajo puede ejecutarse satisfactoriamente con los medios disponibles en la ULPGC y, en su caso, las organizaciones colaboradoras, sin que la EII deba realizar gastos extraordinarios.

Firmas (en caso de que no se haya firmado digitalmente):

Firmas (en caso de que no se haya firmado digitalmente):