

Plan TL

Plan de Impulso de las
Tecnologías del Lenguaje



GOBIERNO
DE ESPAÑA

MINISTERIO
DE ECONOMÍA
Y EMPRESA

PLAN DE IMPULSO DE LAS TECNOLOGÍAS DEL LENGUAJE

CONVENIO SEAD-INCIBE

1. Objeto

2. Actuaciones

3. Contratos licitados

Objeto



Ejecución de proyectos para el cruce entre oferta y demanda de productos y servicios de Ciberseguridad.

Actuaciones

1. Relacionadas con la caracterización de empresas del sector de la Ciberseguridad.
2. Relacionadas con la caracterización del mercado del sector de la Ciberseguridad.
3. Relacionadas con productos y servicios del mercado del sector de la Ciberseguridad.
4. Relacionadas con el cruce de oferta y demanda del sector de la Ciberseguridad.
5. Desarrollo de plataformas comunes.
6. Generación de recursos reutilizables.

1. Servicios TIC para la caracterización de empresas en el sector de la Ciberseguridad.
 - a) Determinación del censo de empresas de Ciberseguridad.
 - b) Obtención de una base de datos de empresas españolas del sector.
 - c) Obtención de un registro de dominios de empresas españolas del sector.
 - d) Revisión de las plataformas de contratación.
 - e) Transformación del catálogo de productos y servicios de Ciberseguridad de INCIBE.

2. Servicios de Ciberseguridad para caracterización de empresas en el sector.

a) Caracterización de empresas del sector:

- Validación manual por expertos de CPVs de contratos.
- Revisión manual de registro de marcas para la identificación de categorías de Ciberseguridad.
- Transformación del catálogo de productos y servicios de Ciberseguridad de INCIBE.

2. Servicios de Ciberseguridad para caracterización de empresas en el sector.

b) Caracterización de mercado:

- Revisión manual por expertos de los contratos clasificados como Ciberseguridad según su CPV.
- Revisión y extracción de indicadores de los corpus para el sector de la Ciberseguridad.

2. Servicios de Ciberseguridad para caracterización de empresas en el sector.

c) Caracterización de los productos y los servicios:

- Selección y validación de los códigos de la clasificación NIZA.



3. Servicios de desarrollo de un sistema de crawling para la caracterización de empresas en el sector de la Ciberseguridad.

- a) Desarrollo del crawler para la identificación de empresas españolas con dominio .es o dominio .com en el sector.

- b) Desarrollo del crawler para el análisis de plataformas de contratación española, europea y alguna latinoamericana.

- c) Procesamiento básico del texto recogido.



4. Ontología de Ciberseguridad para la caracterización del Sector (en coordinación del contrato número 2).

- a) Análisis de la ontología actual, clasificaciones candidatas, productos y servicios.

- b) Diseño de redes ontológicas.

- c) Implementación de la ontología en lenguajes propuestos por el W3C.

- d) Evaluación, documentación y publicación de la ontología.

4. Ontología de Ciberseguridad para la caracterización del Sector (en coordinación del contrato número 2).
 - e) Implementación de un demostrador que permita realizar consultas en SPARQL.
 - f) Extracción de términos de Ciberseguridad a partir de *corpus* disponibles.
 - g) Validación automática e introducción a la ontología.



5. Servicios de Machine Learning para la caracterización de empresas en el sector de la Ciberseguridad.

a) Caracterización de mercado:

- Entrenar motor de machine learning con los CCPPVV validos.
- Lanzar un algoritmo de machine learning para la identificación automática de contratos mal clasificados.
- Lanzar un algoritmo de machine learning para identificar CCPPVV no identificados previamente.

5. Servicios de Machine Learning para la caracterización de empresas en el sector de la Ciberseguridad.

b) Caracterización de empresas:

- Entrenar un algoritmo de machine learning para identificar nuevas empresas en el sector.

- Entrenar un modelo de perfilar empresas en el sector.



5. Servicios de Machine Learning para la caracterización de empresas en el sector de la Ciberseguridad.

c) Para el cruce:

- Entrenar un algoritmo de machine learning para identificar las relaciones entre IIPPCC y patentes.

- Entrenar un algoritmo de machine learning para calcular distancias semánticas en las patentes en el sector.



5. Servicios de Machine Learning para la caracterización de empresas en el sector de la Ciberseguridad.

d) Para calcular modelo de tópicos, en el cual se entrenará un algoritmo de machine learning para desarrollar un análisis de tópicos para el ámbitos de la Ciberseguridad.



6. Distancias semánticas y grafos para la caracterización de empresas del sector de Ciberseguridad.

- a) Cruce textual y semántico para el ámbito de la Ciberseguridad.
- b) Desarrollar análisis de tópicos para el ámbito de la Ciberseguridad.
- c) Entrenar un algoritmo para la identificación entre diferentes fuentes de Ciberseguridad.

Gracias

<https://www.PlanTL.gob.es>

PlanTecnologiasLenguaje@mineco.es