



Jornada sobre Tecnologías de la Traducción para la Administración Pública

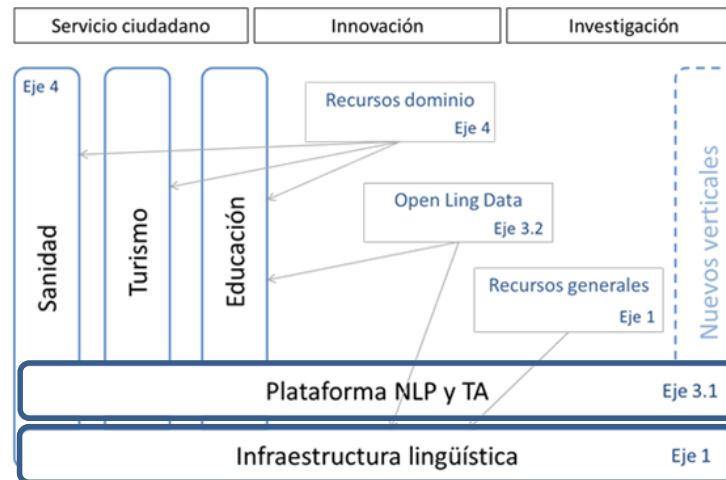
Plan de impulso de Tecnologías del Lenguaje

David Pérez Fernández

Procesamiento de Lenguaje Natural, Sistemas Conversacionales y Traducción automática.

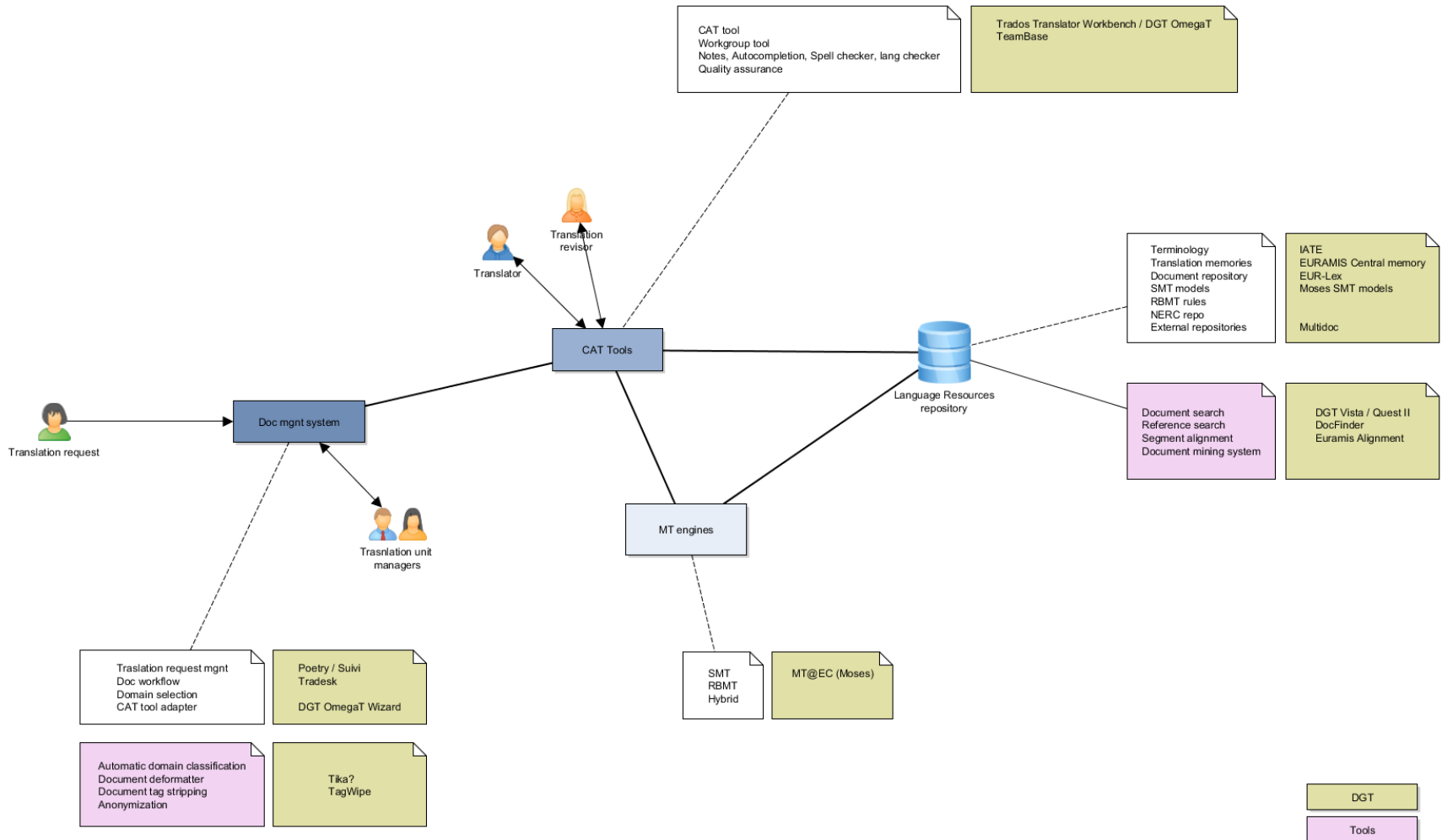
El Plan se articula en los siguientes ejes:

- Eje 1: *Desarrollo de **infraestructuras lingüísticas**.*
- Eje 2: *Impulso de la industria de las tecnologías del lenguaje.*
- Eje 3: *La **Administración Pública** como impulsora de la Industria del Lenguaje.*
- Eje 4: *Proyectos Faro.*



Entorno traducción asistida: Pasado

Computer Aided Translation Environment





ELRC

www.lr-coordination.eu

Gestiona, mantiene y coordina los **recursos lingüísticos** de todos los idiomas oficiales de la UE y países asociados al CEF. Adquisición de recursos lingüísticos adicionales y servicios de procesamiento lingüístico asociados

Suministro al repositorio de recursos lingüísticos de la plataforma CEF eTranslation.



ELRI

www.elri-project.eu

Recopilar, preparar y compartir recursos para la traducción. Red de estaciones nacionales de recursos.

Herramientas extracción recursos y control compartición.

1.6M €, Finaliza OCT 2020

(France, Ireland, Portugal, Spain)



NEC TM DATA

www.nec-tm.eu

Organiza activos nacionales bilingües no explotados que puedan utilizarse como datos, a fin de reducir los costes de traducción en los Estados miembros.

1.3M € Finaliza FEB 2020

(Latvia, Croatia, Spain)



MAPA (Multilingual Anonymisation toolkit for Public Administrations)

ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2019-eu-ia-0013

Herramientas de Procesamiento del Lenguaje Natural para la **anonimización** efectiva y fiable de textos en los ámbitos médico y legal [deidentificación de datos personales].

1.2M €, Finaliza DIC 2021

(Spain, Latvia, Malta, France)



PARACRAWL I y I

www.paracrawl.eu

Generación de **corpus paralelos** para los idiomas europeos a partir de datos descargados de la **web** o colecciones de ficheros PDF.

1.3M €, Finaliza SEPT 2020

(United Kingdom, Spain, Netherlands)



EuroPat (Unleashing European Patent Translations)

ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2018-eu-ia-0061

Extraer **corpus paralelos** de las **patentes** mediante agregación, alineación y conversión de datos de las patentes. Selección de dominio y alta calidad.

927K €, Finaliza SEPT 2020

(United Kingdom, Spain, Netherlands)



NTEU

www.nteu.eu

Construir una **granja de motores** de traducción basados en redes **neuronales**, para idiomas con pocos recursos. Integrado entre los motores de eTranslate.

2.2M €, Finaliza AGO 2021

(Spain, Latvia, Ireland)



MT4ALL

ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2019-eu-ia-0031

Generar nuevos recursos bilingües a partir de datos monolingües. Mejora de motores de traducción automática existentes, y pares de idiomas/dominios no cubiertos.

625K €, Finaliza DIC 2021

(Spain, Ireland, Portugal, Latvia)



iADAATPA/MT-HUB

ec.europa.eu/inea/en/connecting-europe-facility/cef-telecom/2016-eu-ia-0132

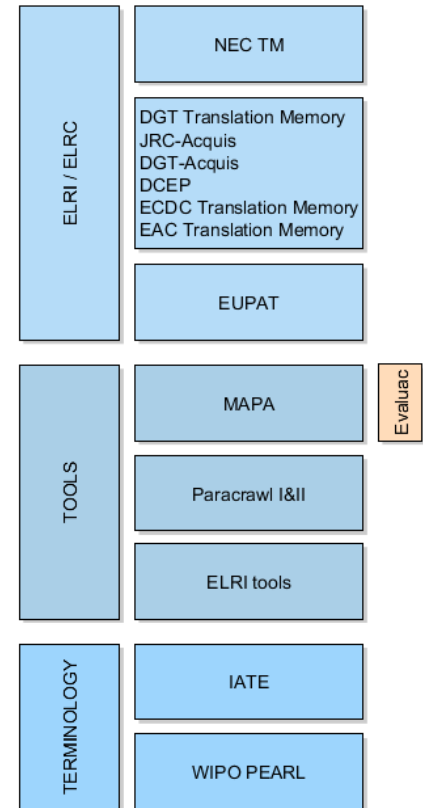
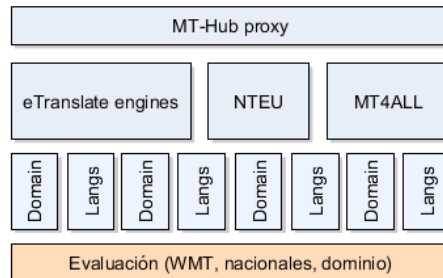
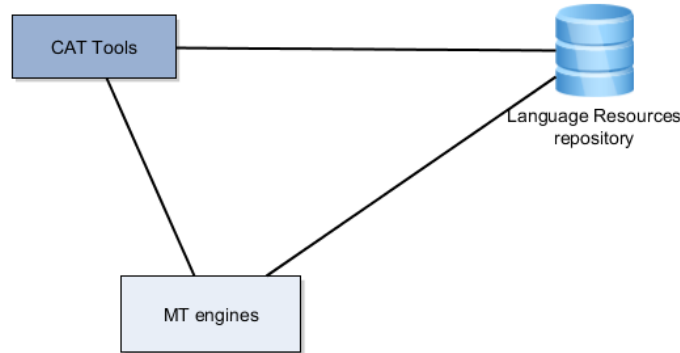
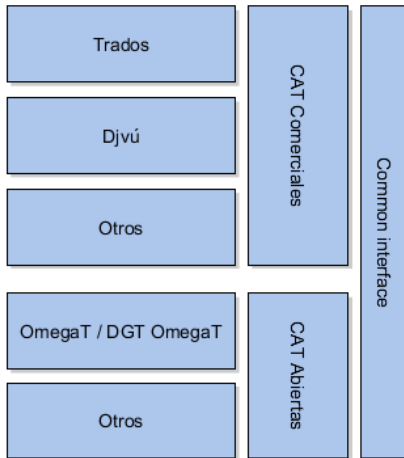
Encaminamiento de servicios de traducción automática (según lengua, dominio, características del servicio), incluye eTranslation CE.

1.9M €, Finaliza FEB 2020

(Ireland, Latvia, Ireland, Spain)


Entorno traducción asistida: Presente


Computer Aided Translation Environment




- Incrementar la **recopilación** de corpus
- Integración en AAPP (ej. SGAD, CCAAs).
Red de nodos de traducción en distintos niveles orgánicos y geográficos
- Mejorar la integración, **estandarización** de interfaces. Resto de DSIs.
- Emplear **campañas** de evaluación (anonimización, traducción ...)
- Más **motores** de traducción: pares de lenguas, dominios, registros
- Mejorar elementos de **despliegue** de sistemas. Integración con entornos para desarrolladores
- Integración sistemas **conversacionales**

¡Muchas gracias!

 www.PlanTL.gob.es

 plantecnologiaslenguaje@mineco.es

 Plan TL - Tecnologías Lenguaje