

## 47. Proyecto PLATA

<b>Nombre del Proyecto</b>	PLATA: Plataforma de Traducción Automática Adaptada a Dominio			
<b>Resumen del proyecto (1 frase)</b>	Desarrollo de una plataforma para la creación y el uso de sistemas de traducción automática adaptados a dominios específicos			
<b>Fecha de comienzo del proyecto</b>	2012	<b>Fecha de fin del proyecto</b>	2017	
<b>Organización líder o coordinadora</b>	Vicomtech-IK4			
<b>Otras organizaciones participantes</b>	<b>Organización</b>	<b>Contribución principal al proyecto</b>		
	MondragónLingua	Explotación de la plataforma		
	MCC Telecom	Integración de la plataforma		
	Universidad Autónoma de Barcelona	Evaluación de calidad y usabilidad de los sistemas adaptados a dominio		
	EITB	Usuario		
	Orona	Usuario		
	Fagor Arrasate	Usuario		
	Alecop	Usuario		
	Galbaian	Usuario		
	Geminys	Usuario		
Danobat Group	Usuario			
<b>Presupuesto del Proyecto (miles euros)</b>	<b>Año</b>	<b>Presupuesto Total</b>	<b>Participación vasca</b>	
	2012	318.757	318.757	
	2013	322.482	322.482	
	2014	316.319	316.319	
	2015	426.218	314.144,5	
	2016	517.947	337.388,4	
	2017	614.545	375.339,3	
<b>Fuentes de financiación de la participación vasca (miles euros)</b>	<b>Año</b>	<b>Financiación 1: Horizonte 2020</b>	<b>Financiación 2: Socios del proyecto</b>	<b>Otras Ayudas públicas</b>
	2012		234.027,65	84.729,35
	2013		230.696,16	91.785,84
	2014		239.008,47	77.310,53
	2015		326.929	99.289
	2016		400.604,25	117.342,75
	2017		475.435,45	139.109,55
<b>Ámbito de actuación</b>	<b>Áreas prioritarias estratégicas</b> <small>Marcar con una X</small>			
	Fabricación Avanzada	Energía	Biosanitaria	
	<b>Territorios de Oportunidad</b> <small>Marcar con una X</small>			
	Alimentación	Hábitat Urbano	Ecosistemas	Ind. Cultural y Creativas
			X	

## **Descripción resumida del Proyecto: principales objetivos y resultados a desarrollar, retos a los que responde, impacto potencial económico y social, etc.**

La traducción automática ofrece ventajas claves para la competitividad de las empresas en sectores claves, al proveer una rápida internacionalización y aumentos de productividad. En particular, la traducción automática es uno de los componentes principales para el crecimiento de un mercado digital único Europeo donde un 43% de los ciudadanos europeos no están dispuestos a adquirir productos o servicios online que no se ofrezcan en su idioma nativo.<sup>1</sup>

Las soluciones de traducción automática son más efectivas para dominios acotados, ofreciendo en estos casos la mayor calidad y usabilidad. Para desarrollar sistemas adaptados a dominios específicos, el paradigma hoy en día dominante es el de la traducción automática estadística, que se basa en los propios datos del dominio. La combinación de estos dos aspectos presenta una oportunidad fuerte para una integración de la tecnología de traducción automática como apoyo al desarrollo de la actividad de empresas en industrias y sectores variados.

El proyecto tiene como objetivo principal el desarrollo de una solución eficiente y flexible para la creación y el uso de sistemas de traducción automática adaptados al dominio específico definido por las actividades de cada empresa. Para lograr este objetivo, es necesario responder a retos tecnológicos en el campo de la traducción automática por un lado, y a retos en términos de usabilidad, integración y uso de la tecnología por otro lado. La meta final del proyecto es un servicio de traducción completamente adaptado a las necesidades de las empresas que ofrezca un apoyo tecnológico óptimo.

### **OBJETIVOS DEL PROYECTO**

Para lograr el objetivo principal, el proyecto debe cumplir una serie de objetivos a varios niveles.

En primer lugar, es necesario desarrollar la investigación en técnicas de adaptación a dominio en traducción automática estadística. Este aspecto cubre la integración óptima de datos de dominios específicos, orientando las traducciones generadas automáticamente hacia los términos y expresiones del dominio. También implica el desarrollo de métodos óptimos para la recopilación y la selección de datos ajenos a un dominio dado, para ofrecer una respuesta adecuada frente a la posible escasez de datos en el dominio considerado. Este aspecto implica el desarrollo de métodos de búsqueda de datos en la red y de combinación de datos variados para lograr conjuntos de datos que permitan el desarrollo de sistemas de traducción óptimos. Por último, es necesario mantener un esfuerzo de investigación fuerte en cuanto a técnicas fundamentales de modelado en traducción automática, para seguir aumentando la calidad y la usabilidad de las traducciones generadas por los sistemas.

En segundo lugar, conviene desarrollar una plataforma que integre los diversos aspectos relacionados con la traducción automática en un entorno único adecuado para el desarrollo eficiente de sistemas de traducción adaptados a diferentes dominios. Al tratar de traducción entre idiomas variados, el campo de la traducción automática cubre una serie de procesamientos que requieren métodos y componentes específicos, cada uno necesario para la creación y el uso de los sistemas finales de traducción. Entre otros, la plataforma necesita integrar componentes de tratamiento de corpus y datos en formatos variados, de modelos de lenguaje y de traducción, de procesos de traducción multilingüe y de gestión de documentos.

<sup>1</sup> [http://ec.europa.eu/public\\_opinion/flash/fl\\_313\\_en.pdf](http://ec.europa.eu/public_opinion/flash/fl_313_en.pdf)

Este último aspecto en particular, permite ganancias de productividad importantes al permitir la generación de documentos traducidos de forma transparente, manteniendo estructura y formato original con contenido textual traducido automáticamente. La plataforma debe proveer una respuesta eficiente a las demandas de creación de sistemas adaptados según las demandas del mercado, y esto incluye la integración de los múltiples componentes necesarios en un entorno configurable que facilite el desarrollo de los sistemas finales.

El tercer objetivo fundamental se sitúa a nivel de usabilidad. La plataforma desarrollada debe responder a las necesidades reales de uso en cuanto a traducciones, necesidades que incluyen varios escenarios: uso directo de traducción automática directa o uso de traducciones automáticas corregidas y validadas por profesionales humanos. Para ello, es esencial dedicar parte del esfuerzo al desarrollo de componentes que permitan un acceso conveniente a las funcionalidades de traducción automática como de post-edición y de validación de las traducciones. La plataforma PLATA incluye así métodos de acceso por servicios web a sus funcionalidades de traducción combinados con un acceso a servicios lingüísticos humanos profesionales, para ofrecer una respuesta flexible según las necesidades de cada empresa.

El último objetivo es el de la evaluación objetiva de las mejoras obtenidas con sistemas adaptados a dominios específicos, comparando la calidad y la usabilidad de las traducciones generadas con sistemas de libre acceso como Google Translate o Microsoft Bing. Al ser basados en datos genéricos del lenguaje, estos sistemas no ofrecen en general una calidad similar que sistemas desarrollados a partir de los datos de dominio específicos y es necesario evaluar precisamente las diferencias de calidad y usabilidad para poder estimar las ganancias de productividad usando sistemas adaptados. Este aspecto permitirá difundir de forma objetiva las ganancias posibles usando la plataforma desarrollada, ofreciendo así datos objetivos para estimar las ventajas posibles adoptando la solución propuesta mediante la plataforma PLATA.

#### **RETOS A LOS QUE RESPONDE EL PROYECTO**

---

El proyecto responde a los retos principales a continuación:

- Generación de sistemas de traducción automática adaptados a diferentes dominios definidos por empresas diversas;
- Optimización de métodos de adaptación a dominio en traducción automática estadística;
- Integración de tecnologías del lenguaje dispersas y fragmentadas;
- Volúmenes de contenido con necesidades de traducción en aumento constante;
- Amplias necesidades de internacionalización que cubren una alta cantidad de idiomas diferentes, en particular en el mercado Europeo;
- Protección de la privacidad de los datos;
- Integración de la traducción automática en procesos empresariales existentes.

## **IMPACTO ESPERADO DEL PROYECTO**

---

Respondiendo a los retos y objetivos mencionados, el impacto esperado del proyecto incluye:

- Aumento de la calidad y usabilidad de las tecnologías de traducción automática;
- Aumento de la integración de tecnologías de traducción automática en procesos empresariales;
- Respuesta más eficiente al aumento de los volúmenes de contenidos con necesidades de traducción;
- Protección de los datos traducidos y mayor confianza en el proceso;
- Reducción de los tiempos de entrada a nuevos mercados para las empresas, en particular las PYMEs;
- Aumento de productividad por parte de las empresas en cuanto a internacionalización y generación de contenidos multilingües y multimodales;
- Respuesta eficiente a las necesidades del mercado digital único Europeo.