



Reciclando ideas

Contenido de este documento

1. Una vision general de lo que hacemos	3
2. Trasfondo del proyecto.....	4
3. Objetivos	4
4. Estado actual de desarrollo.....	5

Índice de figuras

Figura 1: Una visión de conjunto de Patter	3
Figura 2: Pasos de descarga implementados parcialmente	5
Figura 3: Herramienta de alineado (os.align) en fase de test	5

1. Una visión general de lo que hacemos

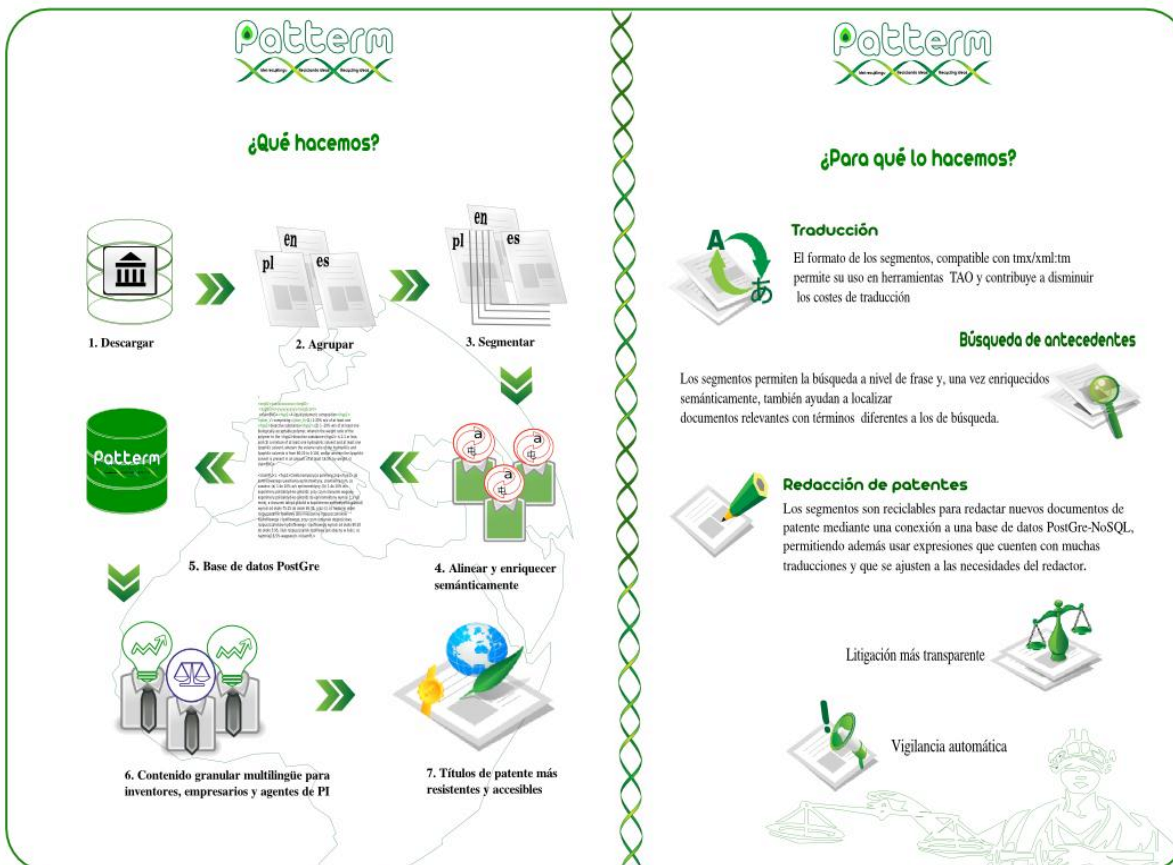


Figura 1: Una visión de conjunto de Patterm

2. Trasfondo del proyecto

Patterm, un juego de palabras con patente (patent), término (term) y pattern (guía, patrón...) es un proyecto de reciclaje de datos, que empezamos hace aprox. 3 años cuando estábamos terminando nuestras licenciaturas en traducción (Gabriel) e informática (Robert).



Robert Mitwicki (Stargard Szczeciński, 1986 (PL)) es programador, padre y especialista en datos de patentes



Gabriel Luis-Franchy (Leverkusen, 1978 (ES)) es traductor, lingüista y especialista en datos de patente.

Creemos en la necesidad de aportar más transparencia y más calidad al sistema de patentes, haciéndolo más asequible a todas las empresas, y no sólo a las grandes multinacionales tal y como ocurre en la actualidad. El sistema de patentes está fuera de control y son muchas las personas que hablan de guerra de patentes o de burbuja de patentes. Un vistazo a los medios como “El País” confirma esta tendencia en todos los ámbitos de la tecnología, especialmente en el ámbito del software.

Veníamos observando el sector de la propiedad industrial y los problemas y costes asociados con ella y pensamos que aplicando tecnología de traducción al mundo de la propiedad industrial podría aportar una serie de ventajas, especialmente para los inventores y empresarios dispuestos a invertir en nuevas tecnologías.

Creemos que las patentes son instrumentos legales muy modernos, pero las fronteras y los diferentes sistemas legales, especialmente fuera de Europa, hacen de ellas un objeto de desprecio, asociado a malas noticias en los medios; litigios entre multinacionales y la lucha de lo privado contra lo público.

Si este es el *status quo* creemos que es posible cambiarlo utilizando la tecnología actual y siempre con los usuarios más importantes en mente: los inventores.

3. Objetivos

Aparte de los objetivos empresariales del proyecto, Patterm tiene como objetivo diseminar la información de patentes multilingües y acercarlas a la comunidad de inventores y lingüistas, así como a la sociedad en general.

4. Estado actual de desarrollo

Pasamos a describir el estado actual de desarrollo guiándonos por la figura 1:



Figura 2: Pasos de descarga implementados parcialmente

Gracias a la colaboración con la Oficina de Polaca de Patentes y Marcas (<http://www.uprp.pl/>) hemos podido automatizar las rutinas de descarga, agrupamiento y segmentación (fraccionamiento del documento en frases o segmentos) de todas las patentes polacas desde 1924 hasta la actualidad. Curiosamente el tamaño en disco original una vez realizada la segmentación es bastante menor de lo que habíamos esperado, lo que nos ha permitido aprovechar mejor los pocos recursos con los que contamos (servidores).

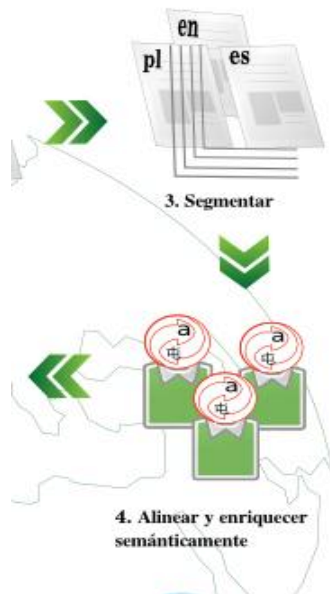


Figura 3: Herramienta de alineado (os.align) en fase de test

La herramienta con la que alinear los documentos no paralelos al 100%, es decir, los que presentan desvíos de contenido a lo largo del ciclo de vida de las patentes, se llama OS.align. Actualmente estamos testando el alineado automático de documentos de patente basado en el uso de expresiones regulares.