



Gestión y producción multilingüe en sistemas de información web: repsol.com, un caso de éxito

Dr. Pedro Luis Díez Orzas
Presidente Ejecutivo de Linguaserve I.S. S.A.

Resumen

Líder en comunicación digital, www.repsol.com ha resuelto sus exigentes y amplias necesidades multilingües gracias a sistemas avanzados. Linguaserve ha contribuido tanto en la interoperabilidad, para el intercambio y gestión multilingüe con el cliente, como en la gestión y producción multilingüe de contenidos, mediante metodologías y herramientas probadas, plataformas integradas y un equipo altamente profesional.

Palabras clave

Web, multilingüe, localización, traducción, *webservice*, XML, Repsol, Linguaserve.

1. Introducción

Suele ser habitual asociar a las grandes empresas y organismos públicos con la utilización de los sistemas más avanzados en materia de Tecnologías de Información. Sin embargo, cuando hablamos de Globalización, Internacionalización, Localización y Traducción (GILT) de Sistemas de Información esto no es siempre cierto.

A menudo, en proyectos de implantación de sistemas de información de cierta envergadura, como un sitio web de gran tamaño y elevado número de servicios, se considera el multilingüismo como una mera traducción, que no requiere atención especial durante las fases de requisitos, especificaciones y desarrollos técnicos. El resultado puede ser un auténtico calvario a la hora de poner en marcha el sistema en varios idiomas y aún más a la hora de mantenerlo en esas mismas lenguas o de añadir una nueva. La sensación que se obtiene no suele ser por tanto muy positiva, percibiendo el multilingüismo en Internet como algo engorroso, caro y difícil de mantener. Así pues, la madurez necesaria para que una empresa dedique recursos expertos y una asesoría bien encaminada al problema del multilingüismo es por lo general el resultado de malas experiencias anteriores.

El caso que mostramos a continuación es en cierta manera la excepción que confirma la regla y el resultado de la experiencia temprana en Internet. Hace más de diez años, Repsol decidió entrar en el medio *online* siendo pioneros en comunicación en Internet. Así, repsol.com nació como una apuesta innovadora y moderna en una plataforma que en aquel momento apenas comenzaba su auge dentro del sector empresarial. Esta plataforma ya era multilingüe y se había adecuado a las necesidades básicas de aquel entonces.

Apenas cinco años después, Repsol se planteó dar un salto tecnológico cualitativo en su web corporativa y las necesidades de multilingüismo se incrementaron y extremaron de forma muy importante. Fue en ese momento crítico cuando supo valorar como un aspecto central el carácter multilingüe de su nuevo portal, por lo que buscó expertos y asesoramiento sobre las mejores soluciones multilingües disponibles, para incorporar después al conjunto del proyecto tecnológico del nuevo portal repsol.com aquella más

conveniente. Linguaserve, empresa especializada en soluciones y servicios multilingües, fue una de las compañías expertas que participaron en aquel proceso, resultando finalmente seleccionada para llevar a cabo el proyecto.

Repsol fue la primera de entre las empresas del IBEX 35 en ofrecer y mantener prácticamente todos los contenidos de su web pública en las lenguas cooficiales, además de otros idiomas. Una de las más grandes webs en volumen de contenidos y servicios, con actualizaciones diarias, más de tres millones de usuarios y múltiples premios nacionales e internacionales, repsol.com es uno de líderes en comunicación digital en España según los más prestigiosos estudios que valoran webs de compañías y referentes en Europa, al estar siempre entre las 10 primeras posiciones.



Fig. 1: Página inicial de repsol.com en catalán (30-06-2010)

2. El desafío

Hoy por hoy, se traduce o interpreta en el mundo menos del uno por ciento de la producción de textos orales y escritos, aunque «el mercado de servicios lingüísticos puede llegar a alcanzar veinticinco mil millones de dólares en 2013» (fuente: *Common Sense Advisory Inc.*). Pero si se tradujera, por ejemplo, un diez por ciento, no habría profesionales suficientes para traducir tanto volumen, ni dinero para costearlo.

Instituciones y empresas han aumentado en los últimos años de forma exponencial sus necesidades de traducción, en términos de número de palabras o páginas. Baste comparar el volumen de traducción que tenía que acometer una empresa hace apenas quince años con lo que requiere hoy día, si desea tener una presencia multilingüe en Internet: de unos miles o pocos cientos de miles a millones de palabras.

También se ha incrementado el número de idiomas a los que se ha de traducir y la terminología que se ha de manejar. Por si fuera poco, la tecnología hace más y más complejo el tratamiento previo y posterior a la traducción en sí, por lo que en muchos casos ésta pasa a ser una actividad más, aunque no menor, integrada en un proceso más amplio: es lo que se denomina localización (L10N) o globalización (G11N). Estos procesos implican, además de la traducción y revisión, de acuerdo con la norma UNE-EN 15038, otras tareas como:

1. Análisis técnico de requisitos y herramientas informáticas.
2. Asesoría técnica, cultural y lingüística.
3. Ingeniería y reingeniería informática.
4. Maquetación y tratamiento gráfico.
5. Adaptación de glosarios terminológicos.
6. Adaptación de traducciones a la interfaz.
7. Adaptación cultural de traducciones.
8. Adaptación audiovisual o multimedia: locución, doblaje y subtitulado.
9. Aseguramiento de calidad lingüístico y funcional.
10. Gestión de proyecto con equipos humanos multidisciplinares.
11. Herramientas de gestión y seguimiento de actividad.

En los sistemas de información multilingües y multiculturales están implicados además requisitos estructurales, funcionales y semánticos; estos requisitos quedan englobados en lo que denominamos "internacionalización" (I18N).

Otra característica de relevancia crucial para la traducción y localización de un sitio web de esta magnitud es la frecuencia y rapidez con que se modifican y actualizan los contenidos, lo que reduce drásticamente los plazos de entrega. La inmediatez de la información en Internet requiere inmediatez en la traducción y localización de esa información, lo que aumenta aún más los costes de la traducción y de la gestión que comporta.

Todos estos factores se dieron cita y confluyeron en el nuevo proyecto de presencia corporativa en Internet de Repsol para afrontar el problema específico del multilingüismo. El desafío pues consistía en crear y mantener un portal que ofreciera de forma continuamente actualizada las versiones idiomáticas en castellano, catalán, euskera y gallego, así como versiones en inglés, portugués de Brasil y portugués de Portugal, además de otras para algunos países de habla hispana. Este desafío debía tener en cuenta además de todos los aspectos anteriores, el hecho de tener un número muy elevado de editores y creadores de contenidos distribuidos en distintos países.

Esto hizo que, antes de lanzar ningún proceso de solicitud de información (*RFI* o *Request for Information*) o de solicitud de propuesta (*RFP* o *Request for Proposal*), procedieran a una fase previa de consultoría y de calibración del alcance tecnológico y traductológico del proyecto, para de esta manera:

1. Valorar los volúmenes de contenidos que se debían acometer, tanto en la fase inicial de migración desde el antiguo portal y de creación de nuevos contenidos, como estimaciones para el posterior mantenimiento.
2. Valorar las metodologías y garantías para obtener un alto nivel de estándar de calidad en los textos traducidos y la dependencia de estos con respecto a la calidad de los textos fuente.
3. Valorar los procedimientos de intercambio, gestión y seguimiento continuo de actualizaciones.
4. Valorar los procedimientos de tratamiento de formatos estructurados y no estructurados, de cara a minimizar las tareas manuales que suelen generar errores y mayores plazos de entrega, con especial

atención a los contenidos en XML con HTML embebido, garantizado el mantenimiento de los niveles de accesibilidad del original (de acuerdo a los estándares W3C).

5. Valorar la creación, explotación y mantenimiento de activos lingüísticos que se generan durante el proceso (fundamentalmente terminología y memorias de traducción): su calidad, volumen y cobertura.

6. Valorar los procedimientos de aseguramiento de calidad tanto de la traducción, como del tratamiento técnico y de la gestión.

7. Y finalmente, aunque sin duda lo más importante, valorar el coste por idioma y el coste global tanto de las fases de puesta en marcha e iniciales como del mantenimiento, definiendo prioridades de acuerdo con el precio, recursos y esfuerzos. Asimismo, calcular el retorno de la inversión en términos corporativos y de acuerdo con los objetivos propuestos.

3. La solución

Entre 2005 y 2006 se acometió la fase de implantación del proyecto, que comprendía migración, traducción inicial y localización de contenidos. Desde entonces, se lleva a cabo un mantenimiento diario de todas las versiones idiomáticas y se han mejorado los sistemas de gestión, seguimiento de trabajos, control y ahorro de costes, así como de los propios procesos de traducción.

Para esto se han aplicado soluciones que hicieran sencilla esta labor y que resolvieran la complejidad y versatilidad del problema. Podemos sintetizar en los siguientes puntos los requisitos básicos y que a nuestro juicio se han cumplido en estas soluciones para las versiones idiomáticas de repsol.com:

1. Las distintas técnicas y metodologías han de estar integradas y no meramente yuxtapuestas. Esto es, deben configurar un único sistema o flujo de trabajo que prevea distintos recorridos en función del tipo de contenido, lengua, formato, prioridad, etc.
2. La tecnología ha de ser un medio y no un fin. Para la consecución con éxito del proyecto, ésta ha de automatizar al máximo lo automatizable y asistir en la medida de lo posible el resto de tareas, de forma segura, robusta y escalable.
3. Las técnicas y herramientas de traducción, terminología y documentación han de utilizarse al límite de sus posibilidades, pero no más. Es mejor asegurar el nivel de calidad que arriesgarlo en aras de hipotéticas reducciones de coste o esfuerzo.
4. Los equipos de trabajo no pueden estar constituidos en compartimentos estancos, sino que ha de existir una coordinación eficaz, una visión global y un cierto conocimiento por parte de unos de las tareas de los otros.
5. El sistema de trabajo ha de ser trazable y reproducible, para resolver incidencias y permitir la automatización del seguimiento por parte del cliente. Han de desarrollarse además no sólo procedimientos, sino también métricas que permitan las mediciones para el establecimiento de mecanismos de mejora.

Lógicamente, para aportar una solución técnica capaz de abordar todos estos factores es necesario partir de plataformas y sistemas previamente concebidos, diseñados y desarrollados, pero que sean al mismo tiempo suficientemente flexibles. Asimismo, han de evitar en la medida de lo posible instalaciones y licencias de software adicionales en el cliente y ser totalmente seguras.

Estas premisas, adoptadas por Linguaserve, se cumplen en el sistema actual de gestión y producción de contenidos multilingües utilizado para repsol.com. Se trata de una potente Plataforma de Localización, Internacionalización y Normalización de la Traducción (PLINT). Para describir esta plataforma y metodología podemos establecer cuatro capas y tipos de actividades dentro del sistema global, que pasamos a describir muy brevemente:

1. La internacionalización previa del sistema de gestión de contenidos del cliente, que requiere al menos las siguientes tareas:
 - 1.1. Conocer el estado previo del sistema de gestión de contenidos.
 - 1.2. Auditoría y en su caso optimización del modelo y organización de datos para la gestión multilingüe.
 - 1.3. Comprobación de *charsets*.
 - 1.4. Auditoría y en su caso optimización del código para el funcionamiento multiidioma.
 - 1.5. Modificación de *workflow* para introducir las nuevas etapas o pasos.
 - 1.6. Auditoría y en su caso modificación de la exportación e importación de contenidos a y de XML.
 - 1.7. Definición y generación de metadatos para la identificación de contenidos que se envíen a través de los *webservices*.
 - 1.8. Valoración de creación y en su caso definición y desarrollo de interfaces de seguimiento.
 - 1.9. Estudio de localización multilingüe de diseño de plantillas.
 - 1.10. Identificación y aprobación de plantillas y elementos funcionales para su adaptación a multiidioma.
 - 1.11. Formatos gráficos y otros formatos: capacidad de gestionar y enviar a traducir archivos fuente de archivos descargables o multimedia.
 - 1.12. Selección y exportación a archivo de contenidos traducibles.
 - 1.13. Accesibilidad (WAI A-AA-AAA).
 - 1.14. Javascript sin cadenas de texto traducible.
 - 1.15. Importación de contenido traducido.
 - 1.16. Plan de pruebas y puesta en producción.
2. Los sistemas de interoperabilidad que se establecen con el cliente son:
 - 2.1. Entre persona y persona (P2P): siempre hay que valorar la posibilidad de establecer métodos de intercambio de información simples, a través de correo electrónico, FTP o incluso soporte físico, para casos específicos o excepcionales.
 - 2.2. Entre sistema y persona (denominado B2C): en este caso, una oficina virtual que permite el intercambio de archivos de forma asistida, el acceso a buscadores potentes para el seguimiento en tiempo real de entregas, costes y estado de los trabajos.
 - 2.3. Entre sistema y sistema (B2B): en este caso, conexión vía *webservices* y formatos estándares de intercambio (SOAP y XML, respectivamente) con los adecuados niveles de seguridad y con el soporte de sistemas que automatizan completamente la selección, codificación, envío, recepción y carga de contenidos, en el lado de Repsol, y la recepción, registro, procesamiento técnico y lingüístico, traducción, revisión, postproducción, entrega y facturación, en el lado de Linguaserve.

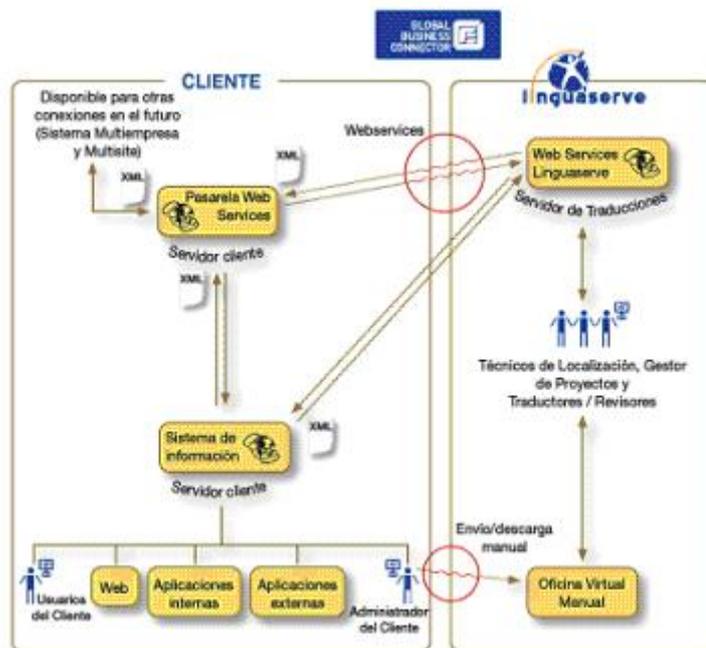


Fig. 2: Proceso general de interoperabilidad

3. El procesamiento y los flujos de trabajo deben contar al menos con los siguientes elementos:
 - 3.1. Capacidad de registrar, consultar y trazar absolutamente todos los pasos del proceso, tanto técnicos como lingüísticos, desde la recepción y activación de un trabajo, hasta la entrega, confirmación, facturación y paso a histórico.
 - 3.2. Mecanismos de control, auditoría y alarmas de los datos registrados.
 - 3.3. Incorporar un sistema totalmente automático para tratamiento técnico de formatos estructurados y asistido para el tratamiento de formatos no estructurados.
 - 3.4. Incorporar sistemas de cooperación con traductores, revisores, terminólogos, técnicos de localización, gestores de proyecto y cliente.
 - 3.5. Utilizar de forma completamente integrada a través de una interfaz de programación de aplicaciones (API o *application programming interface*) herramientas de Traducción Asistida por Ordenador. PLINT incorpora en este momento con magníficos resultados la herramienta Transit XV y está preparado para integrar otras.
 - 3.6. Utilizar controles de calidad durante el proceso: completamente automáticos (sobre formato, por ejemplo) y asistidos (como la terminología o la revisión íntegra de la traducción).
 - 3.7. En los casos en los que sea posible, integrar sistemas de Traducción Automática para que con revisión profesional generen un alto estándar de calidad, totalmente equiparable a la traducción y revisión profesional, que reduzca costes y agilice el trabajo. En este momento se tiene una experiencia probada muy positiva entre lenguas como castellano, gallego y catalán; con portugués, tanto de Portugal como de Brasil, se está en un avanzado estado experimental; y se están empezando a hacer progresos para que en no mucho tiempo sea posible también con el euskera, el inglés, el francés o el italiano.

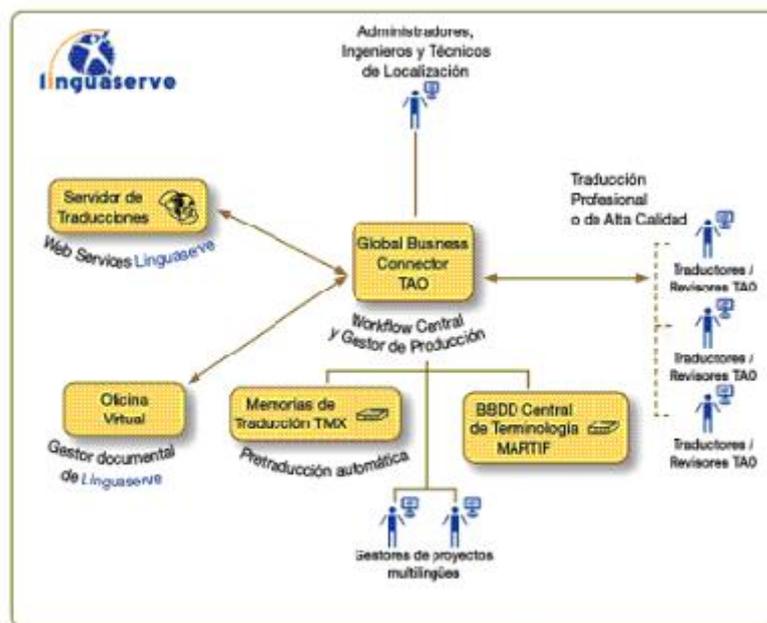


Fig. 3: Plataforma de gestió i producció

4. El almacenamiento de datos y archivos en el servidor de traducción es primordial para el buen funcionamiento de todo el sistema. Ha de ser dinámico, seguro, completo y sobre todo flexible para dar cuenta de la elevada complejidad y casuística existente.
5. La explotación de datos para búsquedas y generación de informes se realiza en tiempo real. Se posibilita así la minimización de tareas de gestión, seguimiento, contabilidad y facturación, así como la toma de decisiones. Todo ello de acuerdo con los sistemas de gestión de calidad de sistema (ISO 9001) y de producto (UNE 15038), así como de seguridad informática (ISO 27001).

La consola de seguimiento tiene acceso controlado por usuario y contraseña con conexión segura, y se puede configurar todo el sistema de envío y recepción de archivos, permitiendo intercambiar cualquier tipo de contenido. Cabe mencionar como ejemplo muy ilustrativo de esta complejidad la guiarepsol.com, parte de repsol.com, que refleja muy bien el papel de empresas GILT, no sólo como traductores y localizadores de contenidos, sino también de servicios y aplicaciones complejos.

El servidor de traducción de Linguaserve ha proporcionado pues todas las características de seguridad para garantizar el acceso exclusivo a los servicios web del cliente y está preparado para recibir todo tipo de contenido por calidad, cantidad y tamaño. Además, minimiza el esfuerzo de configuración y preparación del entorno de trabajo, del tratamiento de los materiales de referencia (como guías de estilo, información de contextualización de los trabajos, terminología y memorias de traducción), de la localización (preproducción y posproducción de formatos y elementos funcionales), del doble circuito de traducción y revisión, de la gestión integral del proyecto y del aseguramiento de calidad, verificación y *testing* del proceso y del resultado final.

Además y por encima de todo lo anterior, la base fundamental para el buen resultado y el alto nivel de satisfacción de los usuarios de repsol.com, auténtico baremo de éxito para Repsol, son los profesionales. La cualificación, formación y experiencia de los profesionales que intervienen en todo el proceso son ciertamente fundamentales: la dirección del proyecto y su sensibilidad para captar las necesidades del cliente, e incluso para anticiparse con una intensa y continua actividad en el campo de la I+D+i; los

gestores de proyecto; los terminólogos; los ingenieros de software e ingenieros y técnicos de localización; y, por supuesto, los traductores y revisores, condición *sine qua non* para el óptimo resultado, que conocen bien herramientas y procedimientos de trabajo, además de ser profesionales contrastados, para lo que aplica el proceso de homologación previsto en la norma UNE 15038.

4. Conclusión

Las Industrias de la Lengua intentan responder a la demanda del multilingüismo en Internet tanto con automatismos para la traducción y revisión, como con soluciones para la gestión, producción y elaboración de materiales de referencia, el tratamiento de los formatos o la interacción entre cliente y proveedor de servicios de traducción (PST). Todos estos procesos, herramientas, sistemas y equipos de profesionales se han de integrar y desplegar con metodologías de trabajo de forma armonizada, para combinar y resolver aspectos y diferencias técnicas y lingüísticas.

En el caso que se ha mostrado, el sistema utilizado por Repsol resulta idóneo para la gestión y producción universal de contenidos (en cualquier idioma, cualquier número de idiomas), así como para el seguimiento de dichas tareas. Capaz de exportar e importar contenidos estructurados y no estructurados, de identificarlos correctamente y de activar el envío y recepción vía *webservices* (bien de forma manual, bien de forma automática), puede tratar los contenidos como documentos, imágenes o multimedia con los archivos fuente para su localización.

Se logra así cubrir con eficiencia todas las necesidades:

- Internacionalización previa del sistema de gestión de contenidos.
- Sistemas de interoperabilidad en su modelo B2C (mediante interfaces entre los usuarios y el servidor de traducción) y sobre todo B2B (mediante interoperabilidad automatizada entre servidores)
- Gestión autónoma y desatendida de los envíos y de las recepciones de los archivos y activación de los flujos de trabajo de traducción y localización.

Son, en última instancia, la satisfacción del usuario sobre el resultado final y la satisfacción del cliente sobre el valor añadido generado durante el proceso los que determinan y constituyen la razón del éxito.