SOPHIA · Caso documentado de QA funcional

Cuando una exportación alteraba la segmentación

Una incidencia real del desarrollo de SOPHIA · febrero de 2026

El mismo contenido, dos resultados diferentes

Durante la integración en SOPHIA de un segmentador de TXT a XLSX, basado en las reglas de segmentación de memoQ, detecté un comportamiento inesperado: el resultado cambiaba según el punto del flujo del que procedía el archivo de entrada.

El TXT normalizado pre-DIFF se segmentaba correctamente. Sin embargo, cuando utilizaba el TXT revisado y exportado post-DIFF, todo el contenido terminaba concentrado en un único segmento.

Comportamiento esperado

9 segmentos

Contenido distribuido en filas independientes, cada una con su correspondiente ID.

	A	B	C
1	ID	Original (DE)	Traducción (ES)
2	1	El buque insignia actual de Sage cafeteras ofrece todo lo que esperamos de una cafetera espresso moderna:	
3	2	Pantalla táctil ultrarrápida	
4	3	14 recetas, 20 ajustes	
5	4	Molinillo Baratza con 45 grados de molienda	
6	5	Tubo de vapor automático (también admite funcionamiento manual)	
7	6	Extracción en frío para cold brew o cold espresso	
8	7	Tiempo de calentamiento breve (30 segundos)	
9	8	Portafiltro profesional de 58 mm + filtros no presurizados	
10	9	Función de cola automática para comenzar a preparar la espuma tras la extracción del expreso	
11			
12			
13			
14			

Captura histórica · febrero de 2026

Comportamiento observado

1 segmento

Todo el contenido concentrado en una única celda del XLSX.

	A	B	C
1	ID	Original (DE)	Traducción (ES)
2	1	El buque insignia actual de Sage cafeteras ofrece todo lo que esperamos de una cafetera espresso moderna:Pantalla táctil ultrarrápida14 recetas, 20 ajustesMolinillo B	

Captura histórica · febrero de 2026

ID	Original (DE)	Traducción (ES)	Notas
1	El buque insignia actual de Sage cafeteras ofrece todo lo que esperamos de una cafetera espresso moderna:Pantalla táctil ultrarrápida14 recetas, 20 ajustesMolinillo Baratza con 45 grados de moliendaTubo de vapor automático (también admite funcionamiento manual)Extracción en frío para cold brew o cold espressoTiempo de calentamiento breve (30 segundos)Portafiltro profesional de 58 mm + filtr		

La pregunta de partida

¿Por qué un segmentador que funcionaba correctamente con el TXT normalizado pre-DIFF dejaba de hacerlo al recibir el archivo exportado después de la revisión?

Localizar dónde se perdía la estructura

Para aislar el problema, comparé los inputs procedentes de distintas fases del mismo recorrido y comprobé los outputs de segmentación.

TXT utilizado como input	Resultado
TXT generado por el Cleaner	**9 segmentos**
TXT de la caja derecha de Translation Prep (normalización pre-DIFF)	**9 segmentos**
TXT guardado antes de editar el DIFF	**9 segmentos**
TXT revisado y exportado desde el DIFF editable (post-DIFF)	**1 segmento**

La comparación permitió localizar la incidencia en la transición de exportación. El archivo conservaba el contenido textual, pero perdía las líneas vacías que el segmentador necesitaba para reconocer las separaciones entre párrafos.

Una función existente ofrecía una pista

Al aplicar Add Blank Lines[2] al TXT exportado desde el DIFF, comprobé que el segmentador volvía a producir los nueve segmentos esperados.

La prueba permitió identificar la condición que provocaba el fallo y verificar una alternativa antes de modificar el flujo.

Sin líneas vacías

TXT exportado post-DIFF

El texto aparece en líneas consecutivas, sin las separaciones necesarias.

```
1  El buque insignia actual de Sage cafeteras ofrece todo lo que esperamos de una cafetera
   espresso moderna:CRLF
2  Pantalla táctil ultrarrápidaCRLF
3  14 recetas, 20 ajustesCRLF
4  Molinillo Baratza con 45 grados de moliendaCRLF
5  Tubo de vapor automático (también admite funcionamiento manual)CRLF
6  Extracción en frío para cold brew o cold espressoCRLF
7  Tiempo de calentamiento breve (30 segundos)CRLF
8  Portafiltro profesional de 58 mm + filtros no presurizadosCRLF
9  Función de cola automática para comenzar a preparar la espuma tras la extracción del
   expreso
```

TXT exportado post-DIFF · febrero de 2026

Con líneas vacías

TXT con separadores recuperados

Se observan las líneas vacías intercaladas mediante los caracteres de fin de línea visibles.

```
1  El buque insignia actual de Sage cafeteras ofrece todo lo que esperamos de una cafetera
   espresso moderna:CRLF
2  CRLF
3  Pantalla táctil ultrarrápidaCRLF
4  CRLF
5  14 recetas, 20 ajustesCRLF
6  CRLF
7  Molinillo Baratza con 45 grados de moliendaCRLF
8  CRLF
9  Tubo de vapor automático (también admite funcionamiento manual)CRLF
10 CRLF
11 Extracción en frío para cold brew o cold espressoCRLF
12 CRLF
13 Tiempo de calentamiento breve (30 segundos)CRLF
14 CRLF
15 Portafiltro profesional de 58 mm + filtros no presurizadosCRLF
16 CRLF
17 Función de cola automática para comenzar a preparar la espuma tras la extracción del
   expresoCRLF
18
```

TXT con separaciones recuperadas · febrero de 2026

[2] Add Blank Lines: función incorporada originalmente a la última pestaña de SOPHIA para añadir separaciones visibles al TXT traducido y facilitar su traslado manual a WordPress. Posteriormente se reutilizó como preprocesamiento opcional antes de segmentar.

Resolver el fallo sin añadir complejidad al recorrido

La primera estrategia fue intervenir en el origen del problema: crear una exportación específica para el modo SOPHIA[1], que conservara tanto los cambios realizados en el DIFF editable como las líneas vacías necesarias para segmentar. La exportación existente para el modo memoQ[1] debía mantener su comportamiento habitual, sin modificaciones.

La implementación necesitó varias iteraciones. La nueva exportación llegó a generar archivos, pero las pruebas revelaron que todavía no producía el resultado esperado. También aparecieron incidencias adicionales durante su integración.

Decidí abandonar ese enfoque y reunir las herramientas standalone en DIY, una pestaña que permite trabajar con memoQ, exclusivamente con SOPHIA o combinar ambos recorridos.

La solución operativa incorporada

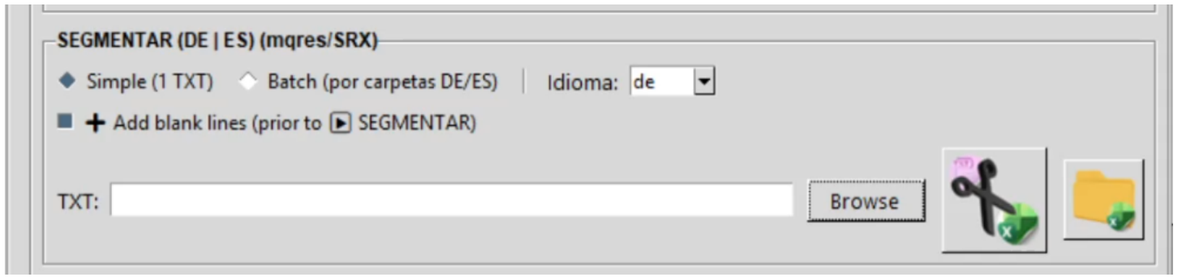
El preprocesamiento Add Blank Lines terminó integrado directamente en DIY, mediante una casilla opcional situada antes de SEGMENTAR y activada por defecto.

De ese modo, podía preparar el TXT en el momento en que lo necesitaba, sin obligarme a pasar por Cierre final para después regresar a una fase anterior.

Qué aportó esta incidencia

- Localizar un defecto mediante la comparación de inputs y outputs en diferentes puntos del proceso.
- Comprobar que una primera modificación no resolvía necesariamente el problema funcional.
- Reconsiderar la vía de implementación para preservar las operaciones existentes y evitar pasos innecesarios.

Implementación posterior



Casilla «Add blank lines» integrada en la función de segmentación de DIY.

srx actual.png · Etapa posterior al intento de Export SOPHIA

Funcionamiento posterior



El vídeo definitivo muestra la transformación de TXT a XLSX con segmentos identificados individualmente.

Fotograma del caso práctico audiovisual de SOPHIA.

Lo que me enseñó esta incidencia

En un sistema integrado como SOPHIA, no basta con comprobar que cada operación funciona de manera aislada. También hay que verificar cómo se relaciona con las fases anteriores y posteriores, y si sus resultados permiten mantener la continuidad del flujo de trabajo.

Cuando una vía de implementación empieza a generar nuevas incidencias sin resolver el problema original, saber descartarla y replantear el enfoque también es una decisión valiosa.

[1] Dos recorridos posibles: el modo SOPHIA permite desarrollar el flujo íntegramente dentro de SOPHIA, sin utilizar memoQ. En el modo memoQ, ambas herramientas se combinan: SOPHIA interviene en determinadas fases y memoQ se utiliza principalmente para la traducción.