

DIFICULTADES Y RETOS EN LA ELABORACIÓN DE PROYECTOS DE HUMANIDADES DIGITALES: EL PROYECTO “MAPPING THE SCRIPTURES IN WESTERN SEPHARDI LITERATURE”*

Soledad Castaño Santos
Universität Basel

RESUMEN: El presente capítulo analiza los retos específicos de la codificación de textos de larga extensión en el marco del proyecto “Mapping the Scriptures in Western Sephardi Literature”, centrado en la edición digital de Biblias romanceadas. En primer lugar, se describen las principales dificultades encontradas en la elaboración de un esquema adaptado a la estructura compleja y a la función litúrgica de este tipo de textos. Se abordan problemáticas como la definición jerárquica de libros, capítulos y versículos, así como la codificación de elementos litúrgicos propios, como las *parashiot* y *haftarot*. En segundo lugar, se analizan los principales retos a los que pueden enfrentarse otros investigadores al trabajar con textos de características no convencionales. A partir de la experiencia práctica acumulada, se proponen soluciones metodológicas que buscan combinar la fidelidad documental con la interoperabilidad de los datos.

PALABRAS CLAVE: Biblias romanceadas, sefardíes, edición digital, TEI, Humanidades Digitales.

Difficulties and Challenges in Developing Digital Humanities Projects: The “Mapping the Scriptures in Western Sephardi Literature” Project

ABSTRACT: This chapter analyzes the specific challenges involved in the encoding of long-form texts within the framework of the “Mapping the Scriptures in Western Sephardi Literature” project, which focuses on the digital edition of Ro-

* Este trabajo se enmarca en el proyecto “Mapping the Scriptures in Western Sephardi Literature” financiando por el SNF (Swiss National Foundation) con el número 204389 desarrollado en la Universität Basel.

mance-language Bibles. First, it describes the main difficulties encountered in the development of a schema adapted to the complex structure and liturgical function of these types of texts. Issues such as the hierarchical definition of books, chapters, and verses, as well as the encoding of distinctive liturgical elements such as *parashiot* and *haftarot*, are addressed. Second, the chapter examines the principal challenges that researchers may face when working with texts featuring unconventional characteristics. Based on the practical experience accumulated during the project, methodological solutions are proposed that aim to combine documentary fidelity with data interoperability.

KEYWORDS: Romance Bibles, Sephardic, digital edition, TEI, Digital Humanities.

1. Introducción

La diáspora sefardí occidental, concebida como tal desde mediados del siglo XVI, constituye uno de los procesos de interés cultural y de evolución identitaria más significativos de la historia europea moderna, con especial énfasis para el ámbito hispano-portugués, lugar de origen de la mayoría de sefardíes exiliados. Este fenómeno ha generado abundante producción académica, concretamente, en el análisis social, religioso y cultural como la reintegración de conversos al judaísmo en las nuevas comunidades europeas. Estudios fundamentales que se han publicado sobre esta temática son los de Yosef Kaplan (1999, 2000), centrados en la identidad y la sociabilidad de las comunidades o los estudios de Harm den Boer (1995, 2003), enfocados a la producción literaria y el ámbito cultural de la comunidad. Estas son obras esenciales para comprender la complejidad de este proceso histórico.

Pese a que diversos estudios han señalado la importancia de las traducciones vernáculas de textos sagrados promovidas desde las comunidades sefardíes europeas, por ejemplo, los estudios sobre producción de traducciones bíblicas de Margherita Morreale (1994, 2007), no existe un estudio exhaustivo sobre ellas. Estas investigaciones esclarecerían la importancia de las obras en la construcción identitaria de la comunidad, especialmente en ámbitos religiosos o educativos. En este caso, es evidente la interrelación entre la producción literaria sacra y la sociedad. Así lo demuestran estudios como el de Ruth Fine (2011) sobre la vida de Daniel Israel López Laguna, poeta y traductor sefardí, y su publicación de *Los salmos de David*, donde se vislumbra el testimonio del exilio.

Un rasgo relevante de estos judíos ibéricos fue el preponderante lugar que ocupó la Biblia en su vida educativa, religiosa y cultural. De este modo, no solo los sefardíes leían y aprendían a través de estas traducciones vernáculas, sino también reconstruían su identidad dotando a las escrituras un importante rol social. Esto les permitía establecer una particular forma de relación con las comunidades cristianas en sus lugares de residencia europeos: Ámsterdam, Hamburgo o Livorno.

La preservación de las ediciones de estos textos sacros es de gran interés para la comunidad judía e, indudablemente, para la comunidad científica. Una gran parte de libros promovidos e impresos en las comunidades sefardíes occidentales se encuentran accesibles en diversas colecciones digitales e instituciones internacionales. La conservación de este corpus constituye una valiosa oportunidad para el estudio de la vida religiosa y política de la comunidad sefardí en la Europa occidental.

Teniendo en cuenta estas circunstancias —la conservación y accesibilidad de las obras, así como la importancia de las mismas en la educación religiosa y social de las comunidades—, se gestó la idea de desarrollar un proyecto que cubriese una investigación sobre la función y el uso específico de los textos sacros entre los judíos conversos de origen ibérico pertenecientes a la diáspora sefardí occidental.

El proyecto se tituló “Mapping the Scriptures in Western Sephardi Literature”, en referencia al corpus escogido y su ubicación espacio-temporal en la concepción histórica. El empleo simbólico de *mapping* ilustra la voluntad de sistematización de las obras existentes con ejemplares conservados y las descripciones bibliográficas. En la actualidad conservamos repertorios sobre este material cultural, entre los más importantes destacan los de Dalen (1690), Rodríguez de Castro (1781) y Kayserling (1859, 1890). Pese a esto, no poseemos una base de datos bibliográfica abierta y especializada con una investigación exhaustiva, de modo que se convierte en una de las vertientes de nuestro proyecto. Por otro lado, seleccionamos una serie de textos significativos y, a partir de ellos, realizaremos investigaciones detalladas sobre aspectos concretos como la funcionalidad del libro en la vida litúrgica, su utilidad en la educación de la comunidad y la construcción identitaria.

Una de las bases fundamentales que articula el proyecto es el compromiso por la incorporación de nuevas tecnologías del campo de las humanidades digitales. Las herramientas digitales nos ofrecen la oportunidad de crear una base de datos en línea y accesible a toda la comunidad científica, partiendo del

repertorio publicado por Harm den Boer (2005), actualizado y ampliado con nuevas investigaciones. Asimismo, el proyecto tiene prevista la publicación de las ediciones digitales de las Biblias romanceadas y otros textos sacros, como *Humashim*, Salterios y Órdenes de oración, con una estructura codificada en XML-TEI.

El objetivo principal es materializar digitalmente los esfuerzos de conservación y preservación del patrimonio, enfatizar el gran valor cultural de este corpus y ponerlo a disposición tanto de la comunidad científica como del público en general, de aquí nuestro interés por el acceso abierto.

Esclarecidos los objetivos de nuestro estudio, corresponde establecer una conceptualización precisa de términos vinculados a nuestro corpus y delimitar el contexto editorial en el que se inscriben, para poder comprender las problemáticas y los retos metodológicos que afrontamos en el desarrollo de la investigación.

La tradición de la Biblia romanceada constituye un fenómeno concreto dentro de la historia de la traducción bíblica en lengua vernácula, a diferencia de las biblias surgidas en contextos protestantes con propósitos reformistas, como la de Reina-Valera (1569), también denominada Biblia del Oso. Las Biblias romanceadas son traducidas exclusivamente por y para las comunidades sefardíes. Estas traducciones poseían una importancia clave en la educación, la religión y la reconstrucción identitaria de una comunidad que no podía disponer ni utilizar la Torah ni otros textos escritos en hebreo porque no conocían la lengua a causa de la persecución inquisitorial. En el contexto de reeducación religiosa, los judíos conversos encontraron un modelo transitorio en estas traducciones para regresar a la fe. Es lógico, pues, que el término “romanceado” que acompaña a estas Biblias haga referencia a las versiones trasladadas del hebreo a lenguas romances como el castellano y el portugués, adaptadas ortográficamente y a nivel léxico a cada comunidad.

El éxito de este fenómeno no habría sido posible sin la labor de los impresores judíos, actores clave en la preservación del patrimonio sefardí, la adaptación del material bibliográfico con fines educativos y la difusión cultural. En la diáspora sefardí occidental, ciudades como Ámsterdam, Livorno, Hamburgo, Londres y Venecia se erigieron como centros neurálgicos de esta actividad. El papel de los impresores ha sido ampliamente estudiado por especialistas como Fuks (1971) y Fuks-Mansfield (1984), Hill (1995), Heller (2007), Van Vliet (2019), Dunkelgrün (2020), y en trabajos más específicos por Heller (1998), Studemund-Hálevy (2002) o Harm den Boer (2005).

Entre todas estas ciudades, Ámsterdam se consolidó como epicentro de la producción bíblica sefardí. La tolerancia religiosa de esta ciudad atrajo a conversos que retornaban al judaísmo, deseosos de recuperar sus tradiciones y prácticas religiosas. El auge de imprentas judías —como las de Menasseh ben Israel y David de Castro Tartas— impulsó no solo la circulación de textos sacros en español, portugués o formatos bilingües (hebreo-romance), sino también la difusión de una cultura escrita que fortaleció la cohesión comunitaria, tal como analizan Kaplan (2000) y Swetschinski (2004). Figuras como Menasseh ben Israel, primer impresor judío en la diáspora, trascendieron el ámbito editorial, actuando como mediador intercultural entre la comunidad sefardí y las sociedades europeas del momento, por ejemplo, la readmisión de judíos en Inglaterra en el siglo XVII.

El auge editorial sefardí surgió como respuesta, en gran parte, a la prohibición de textos sagrados vernáculos en la península ibérica. Tras el establecimiento de la Inquisición, los reinos hispánicos censuraron aquellas traducciones bíblicas no aprobadas por la Iglesia, fenómeno que ha sido analizado por Roth (2002) y Kerner (2018). De este modo, la impresión de Biblias romanceadas se transformó en un símbolo de identidad.

Este fenómeno editorial amsterdamés solo puede analizarse al completo en diálogo con otros factores contemporáneos: el humanismo cristiano y el renacimiento de la imprenta hebrea. El ideal humanista —que propugnaba el retorno a los textos en lenguas originales—, sumado a las innovaciones tipográficas, generó en estas imprentas un ecosistema perfecto para proyectos editoriales que trascendían las producciones en latín. Los sefardíes de la comunidad de Ámsterdam adoptaron un papel activo impulsando el estudio filológico del hebreo bíblico a través de las traducciones al español y el portugués.

Estos aspectos identitarios y filológicos de las Biblias romanceadas plantean, no obstante, desafíos metodológicos al enfocarlos desde las humanidades digitales. El presente artículo ofrece una reflexión crítica sobre las dificultades surgidas en el proceso de selección, transcripción y codificación de estos textos —particularmente por la complejidad lingüística y estructuras multivariantes—, con dos objetivos fundamentales: 1. Enumerar una serie de problemas recurrentes que surgen en el modelado de datos; 2. Compartir las soluciones provisionales implementadas a lo largo del proyecto. Esta perspectiva reflexiva aspira a constituir un marco de referencia para investigadores interesados en el trabajo con textos bíblicos sefardíes o de similar dificultad codicológica.

Antes de profundizar en los aspectos metodológicos de nuestra investigación, es necesario mencionar otros ejemplos vinculados a la literatura sacra que han supuesto un avance significativo en el tratamiento digital de textos religiosos. Estas iniciativas representan soluciones metodológicas específicas a diferentes problemáticas similares a las nuestras, aunque con enfoques distintos según la naturaleza del corpus trabajado. El primer ejemplo se encuentra en *Biblija.net*¹, portal promovido por la Sociedad Bíblica de España y de Eslovenia, que ofrece acceso a todos los libros del Antiguo Testamento y el Nuevo Testamento en diferentes lenguas y versiones. Entre ellas se incluyen las ediciones surgidas por iniciativa de las Sociedades Bíblicas Unidas, la versión de Reina Valera de 1995 o la de Lutero de 1545.

Otro proyecto destacado es *Sefaria*², que pone a disposición del público una amplia colección de textos sagrados judíos, principalmente en inglés. La estructura modular y la interconexión entre textos constituyen un modelo de referencia en cuanto a accesibilidad y navegación textual.

A nivel internacional, destaca la aplicación web de *BibleHub*³, que constituye una herramienta de comparación textual de textos sagrados de diferentes tradiciones.

En el ámbito hispánico, remarcamos la iniciativa *Biblia medieval*⁴, que ofrece un corpus digitalizado de traducciones de la Biblia al castellano realizadas durante la Edad Media. Este proyecto presenta dos versiones de cada transcripción: una paleográfica, que respeta grañas originales, y la normalizada, destinada a facilitar la lectura contemporánea, ambas disponibles a través de enlaces diferenciados.

En esta misma línea de experimentación e innovación, nuestro proyecto aspira a convertirse en un espacio de consulta para investigadores y público interesado en las Biblias romanceadas, que constituyen nuestro principal objeto de estudio.

¹ Portal de *Biblija.net*: <<https://www.biblija.net/biblija.cgi?l=es>>.

² Página web del proyecto *Sefaria*: <<https://www.sefaria.org/texts>>.

³ Puede consultarse en: <<https://biblehub.com>>.

⁴ Disponible en: <<https://corpus.bibliamedieval.es>>.

2. Diseño metodológico para la transcripción y codificación de Biblias romanceadas

La organización y estructuración de los pasos a seguir resultaron fundamentales para garantizar un flujo de trabajo eficiente y ordenado en el proyecto. Desde el inicio, se establecieron una serie de fases que permitieron sistematizar las tareas. Este flujo de trabajo se sintetiza en el siguiente diagrama.



Figura 1

Flujo de trabajo del proyecto “Mapping the Scriptures in Western Sephardi Literature”.

Fuente: elaboración propia

2.1. Recolección y selección de las ediciones de las Biblias

La primera tarea de esta primera fase consistió en la selección y recopilación de los diferentes textos sagrados que conformarían nuestro corpus. Para ello, partimos de los datos ya recogidos en la base de datos de Harm den Boer, así como la actualización de la misma durante los últimos años. Nuestro trabajo inicial fue localizar, verificar y actualizar la información disponible, asegurando así la fiabilidad del material para nuestro corpus.

Respecto a los criterios de selección de los documentos, se priorizó, en primer lugar, la disponibilidad de ejemplares con reproducciones que tuviesen imágenes de alta resolución, concretamente aquellas que siguieran el estándar IIF (International Image Interoperability Framework). Todos los ejemplares que disponían de imágenes con esta calidad fueron seleccionados de manera preferente en nuestra transcripción. Esta alta calidad de imágenes facilita la tarea del siguiente proceso, la segmentación y transcripción de los documentos. En los casos en que no existía una versión en IIF, se optó por aquellos ejemplares que tuviesen la mejor calidad posible, evaluando al mismo tiempo la nitidez del texto reproducido y el estado de conservación general.

Otro de los criterios relevantes que seguimos fue la integridad textual. Se prefirieron impresos que se conservan completos o que presentan pérdidas textuales mínimas. La disponibilidad de textos íntegros es necesaria para ase-

gurar la continuidad en la transcripción y permitir el análisis posterior sin lagunas textuales que puedan comprometer la interpretación entre las diferentes biblias.

Las ediciones de la Biblia romanceada seleccionadas para nuestro corpus son las siguientes:

1. Biblia de Ferrara impresa por Jerónimo de Vargas en 1553.
2. Reedición de la Biblia en Ámsterdam impresa por Yom Tob Atias y Abraham Usque en 1611.
3. Reedición de la Biblia en Ámsterdam impresa por Gilles Jost en 1630.
4. Reedición de la Biblia en Ámsterdam impresa por Gilles Jost en 1646.
5. Reedición de la Biblia en Ámsterdam impresa por Joseph Atias en 1661.
6. Reedición de la Biblia en Ámsterdam impresa por Selomoh Aylon en 1726.

2.2. Segmentación y transcripción usando *Transkribus*

La segunda fase consistió en la creación de un espacio colaborativo de trabajo en la plataforma *Transkribus* (herramienta de transcripción automática con la posibilidad de transcripción manual y la creación de modelos automáticos de reconocimiento de texto). No obstante, el proceso de transcripción no se produjo de manera inmediata. Previamente a la transcripción, la asesoría técnica ofrecida por RISE (Research & Infrastructure Support)⁵ nos planteó la necesidad de establecer una serie de criterios homogéneos, guiados a partir de las siguientes preguntas:

- ¿Qué se quiere transcribir de cada documento?
- ¿Qué se pretende mostrar al público como resultado final?
- ¿Qué características específicas presenta cada uno de los textos seleccionados?
- ¿Cómo deben transcribirse los rasgos peculiares —abreviaturas, variantes gráficas—?

⁵ Agradecemos el asesoramiento técnico proporcionado por la Dr. Elena Spadini y el Dr. José Luis Losada Palenzuela.

- ¿Qué tipo de transcripción —paleográfica, normalizada— es la más adecuada para los fines del proyecto?

La respuesta a estas cuestiones permitió fijar unos criterios de transcripción coherentes y cohesionados. El objetivo de transcripción de estas Biblias fue reproducir una versión del texto principal de forma semipaleográfica, respetando la organización de libros, capítulos y versículos, incidiendo en la marcación identificativa estructural de estos elementos. De este modo, se preserva el contenido del texto, pero también la organización material de estos impresos, esencial para futuras investigaciones críticas.

Tras la toma inicial de decisiones, nos centramos en responder a la tercera de las preguntas. Las ediciones de las Biblias que son objeto de transcripción en este proyecto presentan una serie de características comunes que es necesario sintetizar para comprender mejor los retos que plantean. En las siguientes figuras observamos las características compartidas y singulares de las seis Biblias.

En esta figura observamos los elementos preliminares al texto principal bíblico. Aunque no sean elementos de especial relevancia para nuestra codificación, es importante conocer la estructura de todas las ediciones y, sobre todo, proporcionar una comparativa entre las seis ediciones. Asimismo, recogemos otros aspectos como la utilización de orlas decorativas, que pueden ser susceptibles de estudio en otras investigaciones. Pasando a la segunda parte de la comparativa, nos centramos en los cinco primeros libros, los que constituyen la Torah: Génesis, Éxodo, Levítico, Números y Deuteronomio.

Respecto a la figura tercera, la lista de libros de la denominada Torah en el mundo hebreo, y las dos figuras correspondientes a los libros de profetas, observamos similitudes y diferencias entre las ediciones. Principalmente, distinguimos gran número de similitudes entre 1661 y 1726 en las tres figuras, lo que nos permite inferir que existe un mayor número de modificaciones en estas dos reediciones que en las anteriores, que mantenían total o gradualmente las características de la Biblia de Ferrara publicada en 1553.

Estas dos últimas figuras representan la segunda parte de profetas y otros libros que culminan la estructura de la Biblia. A diferencia de las anteriores figuras, comprobamos cómo las modificaciones entre ediciones han sido menores, esto es, se ha respetado el formato estructural del mismo, incluso cuestiones de versificación que se han ido transmitiendo edición tras edición. Una mayor uniformidad en estos criterios nos asegura una mayor facilidad a la hora de establecer criterios de segmentación y transcripción.

Figura 2
 Tabla comparativa de Biblias (I). Fuente: elaboración propia

BIBLIAS	1553	1611	1630	1646	1661	1726
RASGOS						
Portada	✓	✓	✓	✓	✓	
Dedicatoria	✓	✓				
Prólogo al lector	✓	✓ ² (al lector y al curioso lector)	✓	✓	✓	
Tabla: orden y nombres	✓	✓	✓	✓		
Tabla: sumario de capítulos	✓		✓	✓		
Tabla: catálogo de jueces	✓	✓	✓	✓	✓	
Tabla: orden de las <i>aphtarot</i>		✓	✓	✓		
Tabla: órdenes de <i>pesah</i>		✓				
Tabla: repartición de la Biblia				✓	✓	
Tabla: ley de profetas				✓		
Tabla: <i>parasiot</i>					✓	
Tabla: reglas para <i>parasiot</i>					✓	
Tabla: <i>parasiot</i> en todas las fiestas					✓	
Tabla: lectura antes y después del Meldar						
Pie de página al fin de sección de tablas			✓	✓	✓	
Utilización de orlas decorativas				✓	✓	✓

Figura 3
 Tabla comparativa de Biblias (II). Fuente: elaboración propia

	1553	1611	1630	1646	1661	1726
BIBLIAS						
RASGOS						
CINCO LIBROS (GÉNESIS, ÉXODO, LEVÍTICO, NÚMEROS, DEUTERONOMIO)						
Anotaciones manuscritas	✓		✓			
Incluye <i>parasitot</i> distribuidas en los cinco libros	✓	✓	✓	✓	✓	✓
Éxodo: versos diferenciados señalando órdenes	✓	✓	✓	✓ (2 veces, repetidos en capítulo 28 y capítulo 39)		
Pie de página con número de versículo después de cada uno de los libros	✓	✓	✓	✓	✓	✓
Correcciones			✓			
Nombres hebreos de capítulos					✓	✓
Notas marginales					✓	✓
Numeración de versos					✓	✓
Advertencia al lector sobre las notas marginales					✓	✓
Lista con final de cinco libros y cuenta de versículos	✓	✓	✓	✓	✓ (sin cuenta)	
Pie de página sumando los versículos de los cinco libros		✓				

Figura 4
 Tabla comparativa de Biblias (III). Fuente: elaboración propia

	1553	1611	1630	1646	1661	1726
BIBLIAS						
RASGOS						
	PROFETAS (JOSUÉ, JUECES, SAMUEL I, SAMUEL II, REYES I, REYES II, ISAÍAS, JEREMÍAS, EZEQUIEL, HOSEAN, YOEL, HAMOS, OSADÍAH, JONÁS, MICHA, NAHUM, HABABUK, SEPHANIA, HAGGAY, ZACARIAS, MALACHI)					
Pie de página con fin de libro	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Pie de página fin con número de versículos	<input checked="" type="checkbox"/> (en Josué, Jueces, Reyes II, Isaías, Ezequiel)	<input checked="" type="checkbox"/> (en Josué, Reyes II, Isaías, Jeremías, Ezequiel)	<input checked="" type="checkbox"/> (en Josué, Jueces, Reyes II, Isaías, Jeremías, Ezequiel)	<input checked="" type="checkbox"/> (en Josué, Jueces, Reyes II, Isaías, Jeremías, Ezequiel)	<input checked="" type="checkbox"/> (en Josué, Jueces, Samuel II, Reyes II, Isaías, Jeremías, Ezequiel, Malachi)	<input checked="" type="checkbox"/> (en Josué, Jueces, Samuel II, Reyes II, Isaías, Jeremías, Ezequiel, Malachi)
Josué: cap. 12 versificación de Reyes	<input checked="" type="checkbox"/>					
Jueces: cap. 5 versificación canto de Deborah	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Samuel I: cap. 2 versificación		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>			
Samuel II: cap. 1 versificación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Samuel II: cap. 22 versificación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Reyes I: cap. 8 versificación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Reyes II: cap. 22 separación de versos			<input checked="" type="checkbox"/>			
Isaías: cap. 26 separación de versos			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Isaías: cap. 39 versos sueltos	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/> (cap. 38)	<input checked="" type="checkbox"/> (cap. 38)	<input checked="" type="checkbox"/> (cap. 38)		
Habakuk: cap. 3 versificación	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		
Micha: pie de página con mitad de los 12 profetas en versículos	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		

Figura 5
 Tabla comparativa de Biblias (IV). Fuente: elaboración propia

BIBLIAS	1553	1611	1630	1646	1661	1726
RAGSOS						
Incluye <i>aphtharot</i> distribuidas en los diferentes libros de profetas		✓	✓	✓	✓	✓
Indicador de mitad de libro en versículos		✓	✓	✓		
Indicaciones sobre ausencia de texto		✓				
Indicaciones “hasta aquí” el texto			✓	✓	✓	
Notas marginales					✓	✓
Versos sueltos antes del pie de página					✓	✓
Portada profetas postreros						✓
Reglas para rezo						✓

PROFETAS (JOSUÉ, JUECES, SAMUEL I, SAMUEL II, REYES I, REYES II, ISAÍAS, JEREMÍAS, EZEQUIEL, HOSEAN, YOEL, HAMOS, OSADIAH, JONÁS, MICHA, NAHUM, HABABUK, SEPHANIA, HAGGAY, ZACARÍAS, MALACHI)

Figura 6
 Tabla comparativa de Biblias (V). Fuente: elaboración propia

	1553	1611	1630	1646	1661	1726
BIBLIAS						
RASGOS						
PSALTERIO, PROVERBIOS, YIOB						
Versículos	✓	✓	✓	✓	(subtítulos, pero no versículo)	(subtítulos, pero no versículo)
Indicaciones principio de libro	✓	✓	✓	✓	✓	✓
Pie de página con fin de libro	✓	✓	✓	✓	✓	✓
Pie de página indicando número de versículo	✓	✓	✓	✓	✓	✓
Indicación mitad de libro		✓	✓	✓		
Yiob: cap. 3 versificación		✓		✓		✓
Reglas para rezo						✓

DANIEL, EZRA, JEREMÍAS, ANALES, CÁNTICOS (En las Biblias de 1661, 1726 y 1762 los Cánticos siguen a Yjob, pero el resto se presenta al final de la Biblia)	
Indicación de los 5 envoltorios	<input checked="" type="checkbox"/>
Indicación del inicio de libro	<input checked="" type="checkbox"/>
Pie de página con fin de libro	<input checked="" type="checkbox"/>
Pie de página con número de versículo	<input checked="" type="checkbox"/>
Anales: versificación	<input checked="" type="checkbox"/>
Cántico: versificación	<input checked="" type="checkbox"/>
Indicación de la falta de algún fragmento	<input checked="" type="checkbox"/>
Indicación de la mitad de libro	<input checked="" type="checkbox"/>
Indicación de “hasta aquí”	<input checked="" type="checkbox"/>
Reglas para el rezo	<input checked="" type="checkbox"/>

Figura 7
Tabla comparativa de Biblias (VI). Fuente: elaboración propia

BIBLIAS	1553	1611	1630	1646	1661	1726
RAGSOS						
	RUTH, LAMENTACIONES, KOHELETH, ESTER					
Indicación del inicio de libro	✓	✓	✓	✓	✓	✓
Pie de página con fin de libro	✓	✓	✓	✓	✓	✓
Pie de página con número de versículos	✓	✓	✓	✓	✓	✓
Lamentaciones: versificación	✓	✓	✓	✓	✓	
Ester: nombre de varones a dos columnas	✓	✓	✓	✓	✓	✓
Indicación de mitad de libro		✓	✓	✓	✓	
Reglas para rezo						✓
Versos sueltos previo al pie de página final						✓
Colofón	✓		✓	✓	✓ (orla)	

Resumiendo la información de estas figuras, todas las biblias comparten las siguientes características:

1. Estructura general homogénea siguiendo el esquema de libros, capítulos y versículos.
2. Elementos preliminares paratextuales: presencia de portadas, contraportadas, índices y tablas de contenido.
3. Divisiones del texto en fragmentos como la *parasah* —porciones de la Torah de lectura semanal— y *haftharah* —porciones de los profetas o *Tanaj* que se leen en sábado, fiestas y días de ayuno—.
4. Elementos decorativos.
5. Presencia de léxico hebreo.
6. Uso frecuente de abreviaturas.

La última pregunta que nos queda por responder es la de la toma de decisiones ante la transcripción. Antes de iniciar el proceso de transcripción, establecimos unas normas uniformes que garantizaran la consistencia interna del corpus. Estas decisiones fueron tomadas a partir del diálogo y teniendo en cuenta el objetivo de visualización de las futuras ediciones. En la siguiente lista mostramos las normas seguidas para la transcripción:

1. Normalización de v y u —vno/uno; uvestro/vuestro—.
2. Normalización de ç y c.
3. Resolver abreviaturas —eñl/ en el—.
4. Mantener la abreviatura .A. para referirse a Adonay, siguiendo las indicaciones de los prólogos de las Biblias.
5. Mantener distinciones en nombres propios —Jsrael, Ysreal, Israel—.
6. Mantener las formas verbales aglutinadas —loarlosan—.
7. Mantener la distinción entre f y ph.

Consecuentemente, estas normas favorecen la transcripción de forma semipaleográfica, es decir, una versión fidedigna al impreso, pero que introduce pequeñas modificaciones orientadas a mejorar la inteligibilidad del texto. Asimismo, otra de las decisiones que se tomó antes de empezar la transcripción fue la incorporación de una serie de etiquetas textuales que marcaran regiones estructurales del texto, lo que nos ayudaría con la codificación posterior de la edición en XML-TEI. Las etiquetas principales fueron:

1. *RunningTitlePageNumber*: encabezados de página que combinan título y numeración.
2. *MainText*: cuerpo principal del texto.
3. *CatchwordSignature*: reclamos y firmas.

La utilización de estas etiquetas a nivel de segmentación del texto es crucial. Nos permiten conservar solo el texto principal y poder eliminar segmentos de la transcripción que decidimos no incorporar en nuestra edición como las firmas.

En primera instancia, partimos de otros modelos de transcripción como el de Stefano Bazzaco titulado *SpanishRedonda_sXVI-XVII* y, posteriormente, creamos un modelo propio para las ediciones de la Biblia, que pondremos a disposición al público una vez haya finalizado nuestro proyecto. Transcritos estos seis documentos, procedimos a una revisión exhaustiva para corregir errores y asegurar la homogeneidad del trabajo. La revisión previa a la codificación resulta crucial para exportar los datos correctamente y poder procesarlos en el estándar TEI para la edición digital.

2.3. Codificación con TEI a través de Oxygen

Una de las partes más problemáticas en la elaboración de ediciones digitales se vincula con la toma de decisiones previa a la tarea de codificación propiamente dicha. Desde un principio, consideramos que si nuestro material hubiese pertenecido a un género literario convencional —ya fuera narrativo, poético o teatral—, podríamos habernos guiado por las normas y prácticas establecidas en proyectos anteriores. Tal es el caso del trabajo de Rojas Castro en poesía, el corpus de DraCor para textos teatrales o los proyectos asociados a TEI Publisher, todos ellos con estructuras y esquemas de codificación ya consolidados.

Sin embargo, ¿qué sucede con los textos de amplia extensión, como es el caso de nuestro corpus? Todas las problemáticas que pueden presentarse en proyectos que trabajan con una selección de textos más reducida se magnifican considerablemente cuando el volumen textual es mayor y su complejidad estructural se incrementa. En nuestro caso, se suma además la especificidad litúrgica y lingüística de los textos, lo que impone nuevos desafíos tanto técnicos como conceptuales.

En una primera etapa, procedimos a revisar una amplia variedad de proyectos de ediciones digitales, aun cuando no estuvieran directamente vinculados a nuestra selección textual. El objetivo era identificar metodologías, esquemas de codificación y decisiones documentadas que pudieran servirnos de inspiración o referencia.

Como resultado de este proceso inicial, sintetizamos las etapas clave de nuestro flujo de trabajo en distintas etapas, que representan diferentes tareas: estudio y revisión de las características, estructuración inicial, toma de decisiones editoriales, establecimiento de un esquema final, codificación y revisión final.

La primera etapa implica un análisis detallado de la estructura de los documentos y un primer acercamiento a las posibilidades de etiquetado que ofrece el estándar TEI. Para ello, los investigadores deben familiarizarse con el conjunto de elementos y atributos disponibles, así como con su documentación oficial. La guía TEI resulta de gran utilidad en esta fase, ya que clasifica de forma sistemática todas las etiquetas, proporciona explicaciones claras y ofrece ejemplos concretos de aplicación, lo que facilita una comprensión más precisa de las opciones de codificación. Fue en esta primera fase donde nos encargamos de revisar los elementos y atributos a partir de las guías del estándar TEI y recoger en una lista las que consideramos apropiadas para nuestro objetivo.

La segunda y tercera etapa —estructuración y toma de decisiones— están interrelacionadas, ya que la estructuración no tiene sentido si no está precedida por una reflexión informada sobre las decisiones de codificación. Esta parte del proceso es, con diferencia, la más crítica y donde suelen surgir la mayor cantidad de interrogantes. El investigador debe evaluar todas las características de su texto para decidir qué elementos del lenguaje de marcado TEI se ajustan mejor a las estructuras y funciones que se quieren representar. En muchos casos, esta selección no es evidente y requiere el acompañamiento de técnicos profesionales con experiencia en codificación XML-TEI. Afortunadamente, nuestro proyecto cuenta con el apoyo técnico desde la institución RISE y recibimos una retroalimentación continua para prevenir errores en estas fases críticas.

Es esencial que todas estas decisiones se documenten de forma rigurosa. Para ello, proponemos la creación de una tabla de correspondencias en la que se detallen las características específicas del texto y las etiquetas seleccionadas para representarlas, junto con los atributos asociados, en caso de que los

hubiera. Este instrumento no solo garantiza coherencia a lo largo del proyecto, sino que además se convierte en una herramienta fundamental para la generación del esquema personalizado del proyecto.

Una vez decididas todas estas etiquetas, susceptibles siempre de modificaciones posteriores, se elabora un esquema a través del documento ODD (One Document Does it all), esto es, un archivo en formato XML-TEI que permite generar el esquema técnico en distintos lenguajes, además de conservar la documentación interna del proyecto. Tal como indica Susanna Allés-Torrent (2019): “El documento ODD consiste en un archivo XML-TEI, a partir del cual podremos generar el esquema en diferentes lenguajes y recuperar la documentación de nuestro proyecto”.

Una de las herramientas más comunes para la creación de estos documentos es TEI Roma⁶, una aplicación web que permite tanto crear un nuevo esquema desde cero como importar uno ya existente para su modificación. En TEI Roma, el usuario puede introducir los metadatos iniciales del proyecto —como el título, el identificador, el prefijo del esquema, la lengua de la documentación, entre otros— y proceder luego a la personalización del ODD mediante la selección de elementos, clases de atributos, módulos y tipos de datos. Además, la plataforma permite definir nuevos elementos personalizados según las necesidades del proyecto. Ejemplos prácticos de uso de esta herramienta pueden encontrarse, por ejemplo, en el artículo de Iantorno (2016), dedicado a la edición digital del *Cancionero de Gómez Manrique*.

Una vez consensuado y elaborado el ODD, se pasa a las fases de codificación y revisión del texto en su versión final. Durante estas etapas finales, es fundamental conservar un registro detallado de la actividad y asignar siglas identificativas a cada miembro del equipo encargado de la codificación. Esta práctica facilita el seguimiento del trabajo, la trazabilidad de las decisiones y la colaboración eficaz entre los investigadores.

En nuestro caso, el *software* elegido para la codificación ha sido Oxygen XML Editor, por su robustez, compatibilidad con TEI y funcionalidades avanzadas para trabajar con esquemas personalizados. Asimismo, se ha dispuesto un repositorio privado en GitHub que alojará todos los documentos codificados, permitiendo así un trabajo colaborativo, el control de versiones y la trazabilidad de cada modificación realizada.

⁶ Disponible en: <<https://roma.tei-c.org>>.

3. Retos y estrategias en la codificación de Biblias romanceadas

Uno de los principales retos a los que nos enfrentamos desde el inicio del proyecto fue la definición estructural del corpus. En textos bíblicos, la división en libros, capítulos y versículos es fundamental tanto para su organización interna como para la interoperabilidad con otras ediciones digitales.

El esquema elegido para representar esta estructura se basa en una jerarquía clara utilizando elementos TEI convencionales. Cada libro se codifica como un `<div type="book">`, los capítulos como `<div type="chapter">` y los versículos como unidades de demarcación mediante el elemento `<ab type="verse">`. Esta decisión se tomó tras analizar las ventajas de mantener la estructura clásica de referencia bíblica sin introducir un nivel de anidamiento excesivo. Todas estas etiquetas, libro, capítulo, versículo poseen un identificador propio `@xml:id` que se añadió tras una revisión exhaustiva del modelo estructural de cada una de las seis Biblias. Estos identificadores mediante abreviaturas y números permiten una comparación paralela de los versículos bíblicos, uno de los objetivos de nuestras ediciones digitales. En la siguiente figura mostramos un ejemplo de esta jerarquización estructural. En ella, podemos ver las diferentes divisiones mediante `<div>` y sus diversos `@type`, así como el identificador en cada uno. En este caso, observamos cómo tenemos el libro del Génesis, el capítulo primero y dentro de los versículos los dos primeros versículos que tienen un identificador con la abreviatura del libro más el número del capítulo y el número del versículo. La estandarización de este modelo nos facilita la revisión paralela en todos los niveles jerárquicos, pero, sobre todo, en el caso de los versículos.

```
div type="book" xml:id="GEN" ana="GENESIS">
  <div type="chapter" xml:id="GEN.1" n="1">
    <ab type="verse" xml:id="GEN.1.1" n="1"><lb n="N001" /><hi rend="OrnamentalCapLetter">E</hi><n Príncipe crio el dio
    <lb n="N002" />a los cielos y a la tierra.</ab><ab type="verse" xml:id="GEN.1.2" n="2">
    <lb n="N003" />Y la tierra era vana y
    <lb n="N004" />vazia; y escuridad so-
    <lb n="N005" />bre faças de abysmo; y
    <lb n="N006" />y espirito del dio se mo-
    <lb n="N007" />via, sobre faças de las
    <lb n="N008" />aguas.</ab>
```

Figura 8

Ejemplo de etiquetas jerárquicas en la edición de 1553.

Fuente: elaboración propia

3.1. Elementos de etiquetación compleja

Los elementos paratextuales de las Biblias romanceadas representan un universo complejo que supera en variedad y cantidad al de otros textos contemporáneos. Nos hemos encontrado con portadas, dedicatorias, índices, tablas, glosas marginales o salmos numerados. Nuestra prioridad es la codificación del texto principal, de modo que estos elementos no son de obligatoria transcripción.

Otro de los elementos que es de complejo marcado son los *parasah* y *haftarah*. Siguiendo su definición, corresponden a una división interna estructural que tiene una función litúrgica dependiendo del tiempo de lectura. Las primeras alternativas para su codificación correspondían a la utilización jerarquizada de <div>, pero la disposición de estos elementos a lo largo del texto impedía que funcionase exitosamente. El motivo del error es que estos elementos pueden comenzar a principio de un capítulo y terminar al final de otro, lógico es, pues, que la superposición de divisiones diese error en el lenguaje de marcado TEI.

Se barajaron otras dos alternativas, la primera era marcar estos objetos como nota con el etiquetado <note>, pero no era la función exacta de nuestro elemento. Descartada esa opción, la última alternativa consistía en crear una etiqueta nueva y presentarla al Consorcio TEI, pero comprometería la reutilización de nuestro trabajo. Finalmente, hallamos la solución gracias al diálogo continuo con los técnicos de apoyo. La etiqueta escogida fue , que asocia una anotación con un fragmento de texto. Completamos la etiqueta añadiendo los atributos @n para el número, @from y @to para determinar el inicio y el final del *parasah*. Lo mismo sucedía con las *haftarah* y aplicamos el mismo método con éxito. La codificación de este primer grupo fue relativamente fácil, las *parasiot* son cincuenta y cuatro y corresponden a una semanal por cada año. Además, están ordenadas ordinalmente. Sin embargo, los *aphtarot* no se ordenan ordinalmente, y no existe en todas las ediciones de las Biblias una tabla que las ordene. Desgraciadamente, este factor aumenta la complejidad del etiquetado y del proceso de codificación en general. A fin de ilustrar esta solución, si se observa la siguiente figura, se incluye el funcionamiento de la etiqueta.

```

<div type="book" xml:id="GEN" ana="GENESIS">
  <div type="chapter" xml:id="GEN.1" n="1"><span type="parsha" from="#GEN.1.1" to="#GEN.6.8" n="1"/>
  <ab type="verse" xml:id="GEN.1.1" n="1"><lb n="N001" /><hi rend="OrnamentalCapLetter">E</hi> Principio crió el dio
  <lb n="N002" />sa los cielos y a la tierra.</ab><ab type="verse" xml:id="GEN.1.2" n="2">
  <lb n="N003" />Y la tierra era vana y
  <lb n="N004" />vazia: y escuridad so-
  <lb n="N005" />bre façes de abysmo: y
  <lb n="N006" />y espirito del dio se mo-
  <lb n="N007" />via, sobre façes de lasj
  <lb n="N008" />aguas.</ab><ab type="verse" xml:id="GEN.1.3" n="3"> Y dixo el dio
  <lb n="N009" />sea luz: y fue luz.</ab><ab type="verse" xml:id="GEN.1.4" n="4"> Y vido el dio ala luz que
  <lb n="N010" />buen: y aparto el dio entre la luz y entre la
  <lb n="N011" />escuridad.</ab><ab type="verse" xml:id="GEN.1.5" n="5"> Y llamo el dio ala luz día: y a la es-
  <lb n="N012" />curidad llamo noche, y fue tarde y fue maña-
  <lb n="N013" />na día uno.</ab><ab type="verse" xml:id="GEN.1.6" n="6"> Y dixo el dio sea espondidura
  <lb n="N014" />en medio de las aguas: y sea apartante entre
  <lb n="N015" />aguas y aguas.</ab><ab type="verse" xml:id="GEN.1.7" n="7"> Y hizo el dio a la espondidura:
  <lb n="N016" />y aparto entre las aguas que de baxo a la espa-
  <lb n="N017" />didura y entre las aguas que de arriba a la es-
  <lb n="N018" />pandidura y fue assi.</ab><ab type="verse" xml:id="GEN.1.8" n="8"> Y llamo el dio a la espa-
  <lb n="N019" />didura çielos: y fue tarde y fue mañana día se-
  <lb n="N020" />gundo.</ab>

```

Figura 9

Ejemplo de codificación del elemento parsha en la edición de 1553.

Fuente: elaboración propia

Tanto uno como otro emplean su término en inglés por cuestiones de cohesión dentro del esquema ODD en el lenguaje de marcado TEI. La identificación idéntica en cada una de las seis Biblias permite recuperar estadísticamente el número de ellas presentes en los textos. A aquellas que tuviesen un título propiamente dentro de texto se les añade una etiqueta de <head>. No obstante, algunas ediciones, como la de 1553, no poseen esos títulos y simplemente no se incluye esa etiquetación, pero no existe mayor alteración.

Este ejemplo ha supuesto una de las mayores dificultades a nivel estructural dentro de nuestro proyecto y hemos experimentado con múltiples opciones hasta conseguir dar con la solución más apropiada según nuestros objetivos.

3.2. Regularización ortográfica

La ortografía de los textos romanceados fluctúa entre lo fonético, lo arcaico y lo judeoespañol en contacto. Dado que el objetivo del proyecto es crear una edición semipaleográfica, decidimos conservar la ortografía original, salvo aquellas grafías que habíamos señalado en las normas de transcripción. En el resto de casos, se incluyen sus inconsistencias, pero dejando abierta la posibilidad de representar regularizaciones en futuras investigaciones. El mantenimiento de estas variaciones permitirá nuevas investigaciones en cuanto a la evolución lingüística en las diferentes ediciones.

Respecto a este apartado, se planteó en la toma de decisiones la creación de una versión modernizada que recogiera en el lenguaje de marcado TEI las regularizaciones ortográficas mediante el etiquetado <orig>, <reg> —original, regularizado— o las etiquetas conjuntas <app>, <lem> y <rdg> —aparato crítico, lema y variante de cada testimonio—. Estas formas han sido estudiadas y analizadas para el fenómeno de la variación textual por Rojas Castro (2017) y Fernández Riva (2018).

Siguiendo los objetivos principales para la codificación y sopesando la gran cantidad de inversión material que supondría la regularización, decidimos conservar únicamente una transcripción semipaleográfica, que podrá ser la base de una futura versión modernizada.

3.3. *Hibridación lingüística*

Uno de los aspectos más singulares del corpus es su hibridación lingüística: la coexistencia del castellano, el hebreo y, en menor medida, el arameo. En ciertas Biblias, como la de 1762, que descartamos de nuestro corpus por su dificultad, nos enfrentamos a una verdadera estructura bilingüe. En otras, los términos hebreos aparecen de manera aislada como fórmulas o nombres propios. En la toma de decisiones, propusimos el empleo del atributo @xml:lang sobre palabras que aparezcan en otra lengua diferente a la mayoritaria de la edición impresa. En el caso de que existiesen dos o más lenguas, incluiremos un apartado en el <teiHeader> que identifique el uso de cada lengua.

3.4. *Ejemplos de codificación compleja*

En este último apartado, expondremos algunos casos de codificación compleja que han ido confeccionando el esquema de la codificación de estas ediciones digitales. En primer lugar, la ausencia de títulos de capítulos. A nivel estructural, a la hora de codificar empleamos *scripts* con el lenguaje de programación Python que nos permitiese agrupar los libros y capítulos de forma rápida. La ausencia de título provocaba que no se reconociese el capítulo como tal y tuvimos que corregirlo manualmente.

Otra de las dificultades que complicaron el proceso de codificación fue el empleo de la abreviatura .A. para referirse a Adonay. Estructurados los libros

y capítulos, pasamos otro *script* para separar los versículos por el punto final. En gran parte de las ocasiones, los puntos de la abreviatura se confundían con los puntos finales de oración. La corrección de este problema fue manual para asegurar la cohesión de la abreviatura y la correcta separación de los versículos.

El tercer, y último ejemplo, es la aparición de símbolos no gráficos dentro del texto —asteriscos simples y dobles, paréntesis—. Ya en el proceso de transcripción supusieron un problema por la transcripción errónea que provocaban. A pesar de esos errores, se mantuvieron en las transcripciones. De igual modo, en la elaboración y codificación de versículos volvieron a provocar errores que se solucionaron de forma manual. Estos casos parten de la experiencia propia del proyecto y dan cuenta de la importancia de la revisión continua de la toma de decisiones de un trabajo de tal magnitud.

Este recorrido por las decisiones y problemáticas que atraviesa nuestro proyecto muestra que codificar una Biblia romanceada no se reduce a seguir normas técnicas, sino que exige un trabajo interpretativo, filológico y colaborativo. Apostamos por una codificación modular, sustentada en decisiones documentadas, orientada a la interoperabilidad, y que pueda servir de modelo a otros proyectos con características similares.

4. Conclusiones

La codificación de textos de larga extensión y alta complejidad, como es el caso de las Biblias romanceadas del corpus sefardí, constituye un proceso técnico y filológico de gran exigencia. Este trabajo ha requerido una planificación estructural rigurosa, una toma de decisiones fundamentada en experiencias comparadas y una reflexión continua sobre los límites y alternativas del estándar TEI aplicado a materiales con usos litúrgicos, estructuras multilingües y ediciones heterogéneas.

Uno de los principales logros del proyecto, hasta el momento, ha sido la estandarización de una estructura jerárquica uniforme para los textos bíblicos, basada en la división tradicional en libro, capítulo y versículo. Esta organización se implementó mediante una combinación de elementos `<div type="book">`, `<div type="chapter">` y unidades de marcación versicular con `<ab type="verse">` e identificadores únicos mediante el atributo `@xml:id`. Esto ha permitido establecer una referencia precisa entre los distintos testimonios del

corpus. Esta decisión no solo ha facilitado la navegación interna y la revisión paralela entre biblias, sino que sienta una sólida base para una futura visualización comparativa entre ediciones.

En cuanto a los elementos de etiquetación compleja, como los componentes paratextuales, el proyecto adoptó una política de priorización orientada a la integridad textual. Decidimos codificar solo aquellos elementos que tuviesen una función estructural o litúrgica directa en la comprensión del texto, como las portadas, dedicatorias o glosas relevantes. Casos como las divisiones de *parasah* y *haftarah* plantearon desafíos específicos debido a su distribución no jerárquica, lo que impidió el uso de etiquetas convencionales como `<div>`. La solución final —basada en `` con los atributos `@from`, `@to` y `@n`— fue el resultado de una colaboración estrecha entre investigadores y técnicos, y representa un ejemplo de adaptación válida del modelo TEI a estructuras textuales no lineales. La posibilidad de aplicar esta codificación a todos los textos del corpus, incluyendo aquellas ediciones que carecen de títulos internos, confirma su utilidad.

En el plano lingüístico, la decisión de mantener una transcripción semipa-leográfica responde tanto a criterios filológicos como pragmáticos. La conservación de grafías originales permite futuras investigaciones sobre evolución lingüística, contacto entre lenguas y variación ortográfica en la literatura sefardí. No obstante, la estructura actual del corpus permite el desarrollo de una versión modernizada en fases posteriores sin requerir una reestructuración profunda del marcado.

Otro aspecto fundamental ha sido la gestión de la hibridación lingüística, presente en todo el corpus, aunque con grados diversos. Frente a biblias bilingües como la de 1762 —excluida del corpus precisamente por su complejidad—, las ediciones seleccionadas contienen términos hebreos aislados, fórmulas litúrgicas o nombres propios que exigían una codificación específica. El uso del atributo `@xml:lang` tanto en palabras como en bloques textuales, junto con su documentación en el encabezado `<teiHeader>`, asegura la correcta identificación de las lenguas. Los casos de codificación compleja analizados —como la ausencia de títulos de capítulos, el tratamiento de abreviaturas como .A. para Adonay, o la gestión de símbolos gráficos erróneos— han servido como momentos críticos que permitieron ajustar el flujo de trabajo y definir reglas internas de validación. Estos ejemplos evidencian la necesidad de combinar automatización (a través de *scripts* en Python) con una labor manual de revisión filológica que garantice la coherencia del marcado y su

fidelidad al texto original. Esta alternancia entre procedimientos computacionales y decisiones humanas ha sido clave en el desarrollo de la codificación de las Biblias romanceadas.

En conjunto, la experiencia recogida hasta la fecha revela que la codificación de Biblias romanceadas no puede limitarse a una mera aplicación de reglas técnicas. Se trata de una actividad compleja, en la que intervienen decisiones de orden estructural, filológico, lingüístico y documental. Por ello, el equipo ha optado por una estrategia modular: establecer un marco flexible que permita ir implementando mejoras sucesivas sin comprometer la seguridad del trabajo de corpus ya realizado ni su interoperabilidad.

Uno de los aprendizajes más importantes que extraemos del proceso se encuentra en la importancia de la documentación sistemática de cada decisión, desde la estructura básica del texto hasta las excepciones puntuales. Esto ha derivado en la construcción de un esquema ODD específico del proyecto, que no solo genera automáticamente los esquemas necesarios para la validación técnica, sino que también podrá servir como modelo para futuros proyectos similares.

Por último, cabe señalar que este trabajo, cuya finalización está a disposición a finales de 2026, no busca ser un producto cerrado, sino una base abierta y sostenible para futuras investigaciones. En ese sentido, las decisiones tomadas no solo responden a las características particulares de las Biblias romanceadas del corpus, sino que aspiran a ser útiles para otros proyectos enfrentados a la codificación de textos de gran extensión.

En conclusión, este proyecto se inscribe en una línea de investigación que combina humanidades digitales, filología y estudios sefardíes, y lo hace desde una perspectiva crítica, colaborativa y documentada. Al adoptar el estándar TEI y adaptarlo cuidadosamente a las necesidades específicas de los textos, se ha conseguido preservar un patrimonio textual singular y esperamos poder ofrecer nuevas herramientas metodológicas para que sean reutilizadas por futuros investigadores.

Corpus / Referencias bibliográficas

ALLÉS-TORRENT, Susanna (2019): “Introducción a la Text Encoding Initiative”, *TTHub. Text Technologies Hub: Recursos sobre tecnologías del texto y edición digital*. <<https://TTHub.io/aprende/introduccion-a-tei/>> [consultado 15/04/2025].

- BIBLE HUB (s.f.): <<https://biblehub.com>> [consultado 20/04/2025].
- BIBLIJA.NET (s.f.): “Portal bíblico multilingüe”. <<https://www.biblija.net/>> [consultado 20/04/2025].
- BIBLIA MEDIEVAL (s.f.): “Proyecto de digitalización y edición de Biblias medievales”. <<https://www.bibliamedieval.es/>> [consultado 20/04/2025].
- DALEN, Daniel van den (1693): *Catalogus variorum atque insignium in quavis facultate & lingua librorum... Clarissimi doctissimique Rabbi Samuel Abas*, s.l.: Daniel van den Dalen.
- DEN BOER, Harm (1995): *La literatura sefardí de Amsterdam*. Alcalá de Henares: Instituto Internacional de Estudios Sefardíes y Andalucés.
- (2003): *Spanish and Portuguese Printing in the Northern Netherlands, 1584-1825: Descriptive Bibliography*. Leiden: IDC Publishers.
- (2005): “Isaac de Castro, Albert Boumeester and Early Sephardi Printing in Amsterdam”, *Studia Rosenthaliana*, 38/39, pp. 228-246.
- DUNKELGRÜN, Theodor (2020): “Introduction to the Special Issue: The Jewish Bookshop of the World..”, *Studia Rosenthaliana*, 46/1, pp. 7-28, DOI: <<https://doi.org/10.5117/SR2020.1-2.001.DUNK>> [consultado 15/04/2025].
- FERNÁNDEZ RIVA, Gustavo (2018): “La codificación de la variación textual del paradigma impreso al digital”, *Revista de Humanidades Digitales*, 2, pp. 15-32.
- FINE, Ruth (2011): “De La Liturgia al Relato Testimonial: Los Salmos de David de Daniel Israel López Laguna”, *Caliope*, 17(1), pp. 177-197, DOI: <<https://doi.org/10.5325/caliope.17.1.0177>> [consultado 15/04/2025].
- FUKS, Lajb (1971): “Amsterdam: Hebrew Printing”, *European Judaism: A Journal for the New Europe* 5/2, pp. 17-20.
- FUKS, Lajb y FUKS-MANSFELD, Rena G. (1984): *Hebrew Typography in the Northern Netherlands. Part I*. Leiden: Brill.
- HELLER, Marvin. J. (1998): “Moses ben Abraham Avinu and His Printing-Presses”, [revista sin especificar] 31(2), pp. 123-132.
- (2007): *Studies in the Making of the Early Hebrew Book*. Leiden: Brill. DOI: <<https://doi.org/10.1163/ej.9789004157590.i-420>> [consultado 15/04/2025].
- HILL, Brad S. (1995): “A ‘Catalogue of Hebrew Printers’”, *The British Library Journal*, 21(1), pp. 34-65.
- IANTORNO, Andrea (2016): “Creación de un prototipo para la edición digital del ‘Cancionero’ de Gómez Manrique basado en la aplicación de TEL..”, *Signa*, 25, pp. 169-191.
- KAPLAN, Yosef (1999): “An Alternative Path to Modernity: The Sephardi Jews of Amsterdam in Early Modern Times”, en J. Force y D. S. Katz (eds.), *Everything Connects: In Conference with Richard H. Popkin*. Leiden: Brill, pp. 211-240.
- (2000): *An Alternative Path to Modernity: The Sephardi Diaspora in Western Europe*. Leiden: Brill.
- KAYSERLING, Meyer (1859): *Sephardim: Romanische Poesien der Juden in Spanien*. Berlin: Hermann Mendelssohn.
- (1890): *Biblioteca española-portuguesa-judaica...* Berlin: De Gruyter. DOI: <<https://doi.org/10.1515/9783112338360>> [consultado 15/04/2025].
- KERNER, Alex (2018): *Lost in Translation, Found in Transliteration: Books, Censorship...* Leiden: Brill. DOI: <<https://doi.org/10.1163/9789004367050>> [consultado 15/04/2025].

- MORREALE, Margherita (1994): “La Biblia de Ferrara y los romanceamientos medievales: 2Sm 22 y Ps 18”, en Iacob M. Hassán y Ángel Berenguer Amador (eds.), *Introducción a la Biblia de Ferrara*. Sevilla: Fundación José Manuel Lara, pp. 69-140.
- (2007): “De los sustitutos de la Vulgata en el siglo XVI: La Biblia de Santes Pagnino enmendada por Benito Arias Montano”, *Sefarad*, 67(1), pp. 229-236.
- READ-COOP. (s.f.): Transkribus. <<https://transkribus.eu/>> [consultado 20/04/2025].
- RODRÍGUEZ DE CASTRO, José (1781): *Biblioteca Española. Tomo primero...* Madrid: Imprenta Real de la Gazeta.
- ROJAS CASTRO, Antonio (2017): “La edición crítica digital y la codificación TEI...”, *Revista de Humanidades Digitales*, 1, pp. 4-19. <<https://tinyurl.com/y8oa7fcl>> [consultado 15/04/2025].
- ROTH, Norman (ed.) (2002): *Conversos, Inquisition, and the Expulsion of the Jews from Spain*. Madison: University of Wisconsin Press.
- SEFARIA (s.f.): “Biblioteca de textos judíos en línea”. <<https://www.sefaria.org/>> [consultado 15/04/2025].
- STUEMUND-HÁLEVY, Michael (2002): “Codices gentium Samuel de Isaac Abas, coleccionista de libros hamburgués”, en Jaime Contreras, Bernardo J. García García e Ignacio Pulido (eds.), *Familia, religión y negocio...* Madrid: Fundación Carlos de Amberes/Ministerio de Asuntos Exteriores, pp. 287-319.
- SWETSCHINSKI, Daniel M. (2004): *Reluctant Cosmopolitans: The Portuguese Jews of Seventeenth-Century Amsterdam*. Oxford: Littman Library.
- SYNCO SOFT. (s.f.): *oXygen XML Editor*. <<https://www.oxygenxml.com/>> [consultado 15/04/2025].
- TEI CONSORTIUM. (s.f.): *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. <<https://tei-c.org/>> [consultado 20/04/2025].
- VAN VLIET, Rietje (2019): “Print and Public in Europe 1600-1800”, en Simon Eliot y Jonathan Rose (eds.), *A Companion to the History of the Book*. Hoboken: Wiley, pp. 423-436. DOI: <<https://doi.org/10.1002/9781119018193.ch28>> [consultado 20/04/2025].