

Recetas de las Américas: Un estudio de caso del generador de sitios estáticos Hugo para la publicación bilingüe en la Web¹

Los sitios web estáticos son prometedores para ampliar el acceso a la publicación a grupos a menudo marginados y para volver a conectar a los creadores digitales con los procesos y recursos utilizados para publicar su trabajo, pero no están exentos de desafíos, limitaciones y desigualdades. Aunque los sitios web estáticos ofrecen reducciones a los costos financieros y menor utilización de recursos, tienen el precio de una distribución desigual del trabajo, que suele recaer en el desarrollador técnico (Dombrowski 2022). Debido a la desigualdad lingüística en el software, la documentación y los propios lenguajes de programación, este trabajo se amplifica para los desarrolladores para los que el inglés no es su idioma principal.

Este artículo examinará hasta qué punto la publicación web con Hugo, un generador de sitios estáticos, desafía el dominio de la lengua inglesa en la publicación digital en humanidades y hasta qué punto lo perpetúa a través de un estudio de caso sobre la creación del sitio web estático bilingüe de recetas de periódico *Recetas de las Américas*, desarrollado por el US Caribbean & Florida Digital Newspaper Project (USCFDNP). En este artículo se presenta el USCFDNP y el material original de las recetas que aparecen en *Recetas*, se sigue paso a paso el desarrollo del proyecto y se evalúan los métodos y el software utilizados por su adhesión a los principios de computación mínima y su aplicabilidad a otros proyectos web multilingües de humanidades digitales (HD).

El Proyecto de Periódicos Digitales del Caribe y Florida

Lanzado en 2022, *Recetas de las Américas* (en adelante *Recetas*) es un sitio web estático bilingüe que permite a los usuarios interactuar con recetas publicadas originalmente en el *Diario las Américas*, un periódico hispano en español que comenzó a publicarse

1. En aras de la transparencia, las editoras deseamos aclarar que la traducción del inglés al español de este artículo fue realizada utilizando una combinación de traducción automática y correcciones editoriales humanas. Las citas textuales han sido traducidas de la misma manera desde el original en inglés.

en Miami, Florida, en 1953.² Una década del archivo de este periódico fue digitalizada a través de USCFDNP, un proyecto de colaboración entre las Bibliotecas George A. Smathers de la Universidad de Florida y los sistemas bibliotecarios de la Universidad de Puerto Rico-Río Piedras y la Universidad de las Islas Vírgenes (USCFDNP, s.f.).

USCFDNP forma parte del Programa Nacional de Periódicos Digitales (NDNP), financiado por la Fundación Nacional para las Humanidades (NEH). La subvención NEH NDNP proporciona financiación para la digitalización de periódicos históricos de todo Estados Unidos y sus territorios para el acceso público, para las diversas tareas desde la selección de títulos hasta la digitalización, y para los esfuerzos de concienciación y promoción del uso de los contenidos puestos a disposición pública. El objetivo es preservar y aumentar el acceso a la valiosa información histórica de los periódicos en una plataforma central en línea de libre acceso (NDNP 2023).

USCFDNP ha recibido financiación continua de la NEH durante los últimos 10 años y recientemente adquirió fondos adicionales para ejecutar el proyecto hasta septiembre de 2025. Esta financiación nos ha permitido contratar a una coordinadora del proyecto (cuyo salario proviene 100% de la subvención) que supervisa todos los aspectos técnicos y cotidianos del proyecto. Los fondos también proporcionan apoyo para eventos de divulgación y viajes e, indirectamente, para desarrollar proyectos de humanidades digitales como *Recetas*. Además, gracias a la financiación del NEH, hemos tenido la oportunidad de digitalizar más de 30 periódicos de Florida, Puerto Rico y las Islas Vírgenes desde 1836 hasta 1963, y aún quedan más por realizar.

La digitalización del NDNP consiste en escanear microfilmes para crear archivos TIFF y JPEG 2000 que conservan la imagen del periódico real para su visualización, y en utilizar el reconocimiento óptico de caracteres (OCR, Optical Character Recognition) para crear archivos TXT y PDF para cada página escaneada (NDNP 2023). Las imágenes escaneadas y todos los resultados digitales se ponen a disposición del público de forma gratuita a través de Chronicling America (un sitio web desarrollado y gestionado por la Biblioteca del Congreso para el NDNP), así como a través de las Colecciones Digitales de la Universidad de Florida (UFDC) y la Biblioteca Digital del Caribe (dLOC).³ Los resultados digitales exigidos por la subvención permiten que el contenido de los periódicos pueda ser utilizado libremente como datos para HD y otros proyectos por cualquier persona.

2. Sarah Tew y Melissa Jerome, *Recetas de las Américas*, <https://recetas.domains.uflib.ufl.edu>.

3. Library of Congress, Chronicling America, consultado el 31 de diciembre de 2023, <https://www.loc.gov/collections/chronicling-america>; University of Florida George A. Smathers Libraries, University of Florida Digital Collections (UFDC), consultado el 31 de diciembre de 2023, <https://ufdc.ufl.edu>; y Biblioteca Digital del Caribe (dLOC), consultado el 31 de diciembre de 2023, <https://dloc.com/es/>.

Recetas de las Américas

Recetas nació del deseo de incitar a los usuarios a utilizar los datos de los periódicos históricos digitalizados como parte del USCFDNP. Para promover el contenido de los periódicos, buscamos desarrollar un proyecto de humanidades digitales que captara la atención de los usuarios y que, al mismo tiempo, tuviera una barrera técnica baja para que pudiera reproducirse en otros conjuntos de datos de periódicos. La sección “Receta del Día” del *Diario las Américas* demostró ser un buen punto de partida: se publicaba una receta en casi todos los números del periódico y la mayoría de las recetas seguían un formato estándar que incluía el encabezado “Receta del Día”, el título de la receta, la lista de ingredientes y las instrucciones.

La comida también parecía ser una manera fácil de atraer a una gran audiencia, ya que es esencial para nuestra existencia humana y sirve como conector a través de las fronteras sociales y geográficas. Aunque muchos de los otros periódicos que hemos digitalizado a través del USCFDNP incluyen recetas, decidimos centrarnos en las impresas en el *Diario las Américas* para reconocer los lazos pasados y presentes entre las comunidades de Florida, América Latina y el Caribe. Resaltar el contenido de origen hispano también era personalmente importante para nosotras, ya que ambas trabajamos en la Colección Latinoamericana y Caribeña de las Bibliotecas de la Universidad de Florida y Melissa Jerome es latina.

Recetas consta actualmente de 53 recetas publicadas originalmente en español en *Diario las Américas* entre 1954 y 1960. Pueden explorarse de múltiples maneras (Jerome 2022). La pestaña Recetas ofrece una lista (por orden alfabético) de los títulos de las recetas disponibles. Desde esta página, los usuarios también pueden utilizar el cuadro de búsqueda para encontrar una receta por nombre de receta (Figura 1).

Cualquiera de las dos opciones permite a los usuarios hacer clic en la receta que deseen y visualizarla. La pestaña Filtro permite una búsqueda más exhaustiva. Con esta opción de búsqueda, los usuarios pueden elegir entre tres categorías de etiquetas en la parte izquierda: (1) restricciones dietéticas, (2) tipo de plato y (3) ingredientes. Seleccionando cualquier combinación de estas etiquetas se obtiene una lista de recetas que incorporan las restricciones dietéticas, el tipo de comida o los ingredientes seleccionados. Además, los usuarios tienen a su disposición un cuadro de búsqueda en el que pueden buscar una receta por ingrediente (Figura 2).

Debido al pequeño tamaño de la muestra y a la naturaleza de los sitios estáticos, el desarrollo de *Recetas* desde la ideación hasta la creación fue relativamente rápido. Pasamos de tener una visión en julio de 2022, a recopilar, limpiar y traducir los datos, para finalmente lanzar públicamente la primera iteración del sitio tres meses después, en octubre de 2022. Desde su lanzamiento hasta diciembre de 2023, el sitio ha recibido 3082 visitas (2118 visitantes únicos) con 8716 páginas servidas. El sitio ha recibido



Figura 1. Página de lista de recetas en español. Consultada el 15 de abril de 2024.



Figura 2. Página de filtro en español. Consultada el 15 de abril de 2024.

comentarios positivos del público. Por ejemplo, un profesor de otra institución expresó su interés en trabajar con nosotros para incorporar este proyecto a su curso de español. Además, otras personas han probado las recetas y han enviado fotos, que aparecen en la página de inicio del sitio. Este interés y participación ponen de manifiesto la importancia de la accesibilidad y la inclusión en los proyectos de HD.

En el mundo académico y en el campo de las HD, el inglés se ha impuesto como lengua por defecto. En el Norte Global, las probabilidades de encontrar estudios que no estén en inglés son escasas, y a menudo se limitan a ser publicados en números especiales con lenguas designadas, lo que “refuerza aún más el desequilibrio de poder entre las lenguas” (Nilsson-Fernández y Dombrowski 2022, 86). Dado que combatir la hegemonía del inglés es un objetivo personal y profesional para el equipo, y que el contenido de recetas de nuestra web se publicó originalmente en español, *Recetas* se creó como un sitio bilingüe y utiliza el español como idioma por defecto. La publicación en español representa y conecta mejor a los usuarios con los orígenes culturales de las recetas, sus creadores y el público al que van dirigidas, mientras que la inclusión de traducciones al inglés tiende puentes culturales, refleja el pasado y el presente multicultural de Florida y favorece una mayor accesibilidad e involucramiento con los materiales.

Para lograr la funcionalidad bilingüe del sitio, nos apoyamos en gran medida en nuestro dominio de las lenguas habladas. Dominar tanto el español como el inglés resultó ventajoso, ya que nos permitió discernir matices y complejidades contextuales que los procesos de traducción automática pasan por alto. Al principio, recurrimos a la función de traducción automática de Microsoft Excel para agilizar el proceso de traducción. Aunque al principio este método proporcionaba traducciones bastante precisas, requería una revisión exhaustiva de la ortografía, la gramática y la precisión contextual. Aunque se consideró la posibilidad de recurrir a servicios de traducción profesional, el elevado costo asociado y la limitada disponibilidad lingüística presentaron desafíos (Nilsson-Fernández y Dombrowski 2022). Además, las limitaciones presupuestarias nos impidieron subcontratar en esta tarea, por lo que tuvimos que asumirla nosotras mismas.

Computación mínima

La computación mínima reduce las barreras relacionadas con los costos, los conocimientos técnicos y el acceso a hardware y software. Es útil en los casos en que la innovación y la creatividad son esenciales, especialmente cuando hay poca o ninguna financiación disponible para invertir en programas patentados que pueden no permanecer accesibles o disponibles a largo plazo (Risam 2018). La computación mínima también facilita la exploración de la publicación multilingüe, que es particularmente significativa para los profesionales de HD que operan en idiomas distintos del inglés.

El trabajo en HD en el Norte Global normalmente eclipsa el valioso trabajo en lengua no inglesa de otros en el campo (Gil y Ortega 2016). Proporcionar contenido solo en inglés da como resultado una falta de contexto y narrativas ausentes o malinterpretadas que entran en conflicto con los objetivos de las HD de ser interseccionales e inclusivas (Nilsson-Fernández y Dombrowski 2022). Existe la necesidad en el campo de las HD de “superar el sesgo anglocéntrico” para adaptarse a las necesidades multilingües de la creciente comunidad mundial de las HD (Tanasescu 2021, 38). Es crucial que los profesionales de las HD consideren “herramientas, tutoriales y códigos que satisfagan las necesidades de los académicos que trabajan en idiomas distintos del inglés” (Tanasescu 2021, 30). Una herramienta que los profesionales de las HD pueden utilizar para dismantelar estas barreras es la computación mínima. La computación mínima abre las puertas al “desarrollo de un registro cultural digital que incluya voces e historias de comunidades que han sido eludidas en el registro cultural” (Risam y Gil 2022, párr. 21). Esto es lo que impulsa nuestro trabajo con *Recetas* y la razón por la que tomamos ciertas decisiones lingüísticas y técnicas para garantizar que reconocemos y representamos a las comunidades cuyas voces han sido históricamente y son continuamente subrepresentadas y con la intención de crear un modelo que pueda ser compartido y replicado por otros.

La computación mínima se plantea como un “mínimo indispensable para lograr tareas sencillas” (Gil y Ortega 2016, 28), lo que la convierte en una opción atractiva tanto para los que pueden ser nuevos como para los que tienen más experiencia en el campo. Sin embargo, este enfoque no es “mínimo”: hay un trabajo que a menudo se distribuye de forma desigual, recayendo la mayoría de las veces en el desarrollador técnico (Dombrowski 2022). “Mínimo” tampoco significa facilidad de uso, sino más bien “aboga por utilizar *sólo* las tecnologías que son necesarias y suficientes para desarrollar el trabajo en humanidades digitales” (Risam y Gil 2022). Mantenernos centradas en, como dicen Alex Gil y Élika Ortega, “lo que necesitamos” fue el motor de nuestras decisiones de mantener *Recetas* “mínimo” y publicarlo como un sitio estático multilingüe.

La publicación de páginas web estáticas

Hicimos *Recetas* con Hugo, un generador de sitios estáticos, para ayudar a la gente a acceder mejor a las recetas históricas de los periódicos.⁴ Los sitios estáticos son una de las muchas herramientas de computación mínima y muy adecuados para la publicación multilingüe. Optamos por un sitio estático para publicar *Recetas* en lugar de

4. Hugo, consultado el 31 de diciembre de 2023, <https://gohugo.io>.

enfoques más tradicionales que consumen muchos recursos, como WordPress y Drupal, porque un sitio estático requiere menos costos financieros, computacionales, de almacenamiento, de conservación y de mantenimiento. El tamaño del sitio *Recetas* terminado es actualmente de 41 MB, y el tamaño de los materiales utilizados por Hugo para crearlo es de 108 MB. El sitio web se genera en 1030 ms (aproximadamente 1 segundo).

Los sitios estáticos son una buena opción para los proyectos de HD por su bajo costo. Utilizan tecnologías web básicas y son especialmente adecuados para proyectos de contenido fijo que no requieren funciones dinámicas ni actualizaciones frecuentes. Debido a su simplicidad, también cuestan menos tiempo al usuario y requieren menos recursos informáticos para acceder a ellos. Estas características respaldan nuestro objetivo de crear un modelo de publicación más accesible, sobre todo para quienes trabajan fuera de las instituciones tradicionales, con menos recursos financieros y tecnológicos, y/o en idiomas distintos del inglés.

Aunque los generadores de sitios estáticos pueden combinarse con sistemas de gestión de contenidos (CMS, *Content Management System*) sin cabeza, decidimos no añadir un CMS a *Recetas* y confiar únicamente en la estructura de directorios de archivos de Hugo para la gestión de contenidos.⁵ Al eliminar todas las demás dependencias de software, *Recetas* es más pequeño, más ligero y menos complejo técnicamente, al tiempo que se reducen las vulnerabilidades del sitio y, por tanto, se ahorra dinero y tiempo del personal informático.

Hugo es muy similar a otros generadores de sitios estáticos, incluyendo Jekyll, que está detrás de muchas herramientas de uso frecuente en HD como GitHub Pages, Collection Builder y Wax. Los sitios estáticos para proyectos de HD están bien documentados, y los profesionales de HD, incluyendo a Kate Newson (2017) y Chris Diaz (2018) han publicado guías prácticas para entender y usar generadores de sitios estáticos para publicaciones académicas y proyectos digitales.

Aunque los menores costos financieros y computacionales de los sitios web estáticos pueden ampliar potencialmente el acceso a la publicación digital a quienes están fuera de los modelos tradicionales, algunas de estas ventajas se obtienen a costa del tiempo y el trabajo del desarrollador (Dombrowski 2022). En la elaboración de *Recetas*, el tiempo y las habilidades para la creación de plantillas tuvieron el precio más alto.

Los generadores de sitios estáticos necesitan plantillas para compilar los distintos componentes de un sitio. Aunque los principios de la creación de plantillas son transferibles, los distintos generadores utilizan lenguajes diferentes para esta labor, cada uno con su propia sintaxis, variables, funciones y métodos. Las plantillas son intransferibles. Las plantillas y otros recursos pueden agruparse en un tema. El sitio web de Hugo

5. Para una breve visión general de los CMS sin cabeza con sitios estáticos, véase Newson (2017).

ofrece 250 temas gratuitos diferentes, la mayoría para blogs, portafolios digitales y documentación (Hugo, s.f.-a).

Escribir plantillas y entender la jerarquía de plantillas y temas de Hugo fue el aspecto más difícil del proceso de publicación de *Recetas*. Para *Recetas*, Sarah Tew escribió plantillas individualizadas para los distintos tipos de páginas que requería el proyecto y un tema básico para todo el sitio. Los proyectos de HD suelen ser únicos y requieren diseños y capacidades más allá del blog tradicional o del sitio de comercio electrónico básico que las plantillas gratuitas existentes no suelen ofrecer. La página de filtros de *Recetas* es un ejemplo de este problema, que se resolvió escribiendo una plantilla individualizada con JavaScript.

Hugo multilingüe

Elegimos Hugo en lugar de otros generadores de sitios estáticos porque Hugo ofrece un mejor manejo multilingüe. Hugo y muchos otros generadores de sitios estáticos pueden rellenar automáticamente algunos textos básicos, como fechas y horas, en varios idiomas, pero no traducen el contenido. Hugo es una buena opción para proyectos multilingües porque hace que el contenido multilingüe y su organización sean iguales y directamente visibles, no asume ningún idioma como idioma principal del sitio a menos que se indique explícitamente, y no privilegia la publicación monolingüe requiriendo plugins o descargas extra para idiomas adicionales.

A diferencia de los generadores de sitios estáticos como Jekyll y Eleventy, Hugo incluye soporte multilingüe directamente en el paquete Hugo (Hugo, s.f.-b). Una vez descargado e instalado, Hugo puede gestionar un sitio multilingüe siempre que el desarrollador sepa dónde y cómo declarar los idiomas y organizar el contenido. Eleventy, Gatsby y la mayoría de los demás generadores de sitios estáticos dependen de plugins para añadir soporte multilingüe. Esto añade complejidad técnica y más posibilidades de que se rompan los procesos y se produzcan frustraciones, especialmente para los desarrolladores menos experimentados. Hay formas de configurar un sitio Jekyll multilingüe sin plugins, pero estos métodos son más complejos que los de Hugo (Simple IT Rocks 2019). Eleventy también maneja bien el contenido multilingüe, pero con más complejidad, ya que requiere JavaScript para localizar fechas, por ejemplo.

Hugo también tiene tiempos de construcción más rápidos que otros generadores de sitios estáticos. Aunque no es una preocupación para nuestro proyecto en este momento, esperamos ingerir cientos, si no miles, de recetas de otros periódicos en el futuro. A medida que aumenta la escala, la velocidad de construcción se vuelve cada vez más valiosa, especialmente en las pruebas de desarrollo, ya que el sitio se construye

y reconstruye localmente a medida que el desarrollador realiza cambios. Incluso para los proyectos pequeños, el tiempo de compilación puede convertirse en un factor limitante y frustrante a medida que se añaden archivos de imagen de gran tamaño o con zoom profundo. Los sitios web multilingües son más grandes que los monolingües, ya que existe la misma página en cada idioma, por lo que el tiempo de compilación más rápido se convierte en una ventaja mayor cuanto más grandes y más archivos y más idiomas tenga un sitio, lo que convierte a Hugo en una buena opción para proyectos multilingües.

Suponiendo que el desarrollador tenga los conocimientos técnicos para escribir sus propias plantillas, otra ventaja de los sitios estáticos (hechos con Hugo u otros generadores) para proyectos multilingües es la posibilidad de crear plantillas únicas para contenidos en diferentes idiomas. Aunque *Recetas* no explota esta capacidad—las plantillas en español e inglés son las mismas, sólo cambia el contenido que las rellena—es útil para proyectos que requieren diferentes diseños para la lectura, acomodando idiomas de izquierda a derecha, de derecha a izquierda y verticales todos en el mismo sitio web o incluso en la misma página de forma más sencilla que otros métodos de publicación web como WordPress, Drupal y Omeka. Sin embargo, el propio lenguaje de plantillas presenta síntomas de monolingüismo inglés. Go, el lenguaje en el que se basa Hugo, fue creado en Google por tres programadores informáticos de Canadá, Suiza y Estados Unidos con el inglés como idioma aparente por defecto en el código. Aunque Hugo tiene un gran potencial para facilitar la publicación web en idiomas distintos del inglés, no socava el dominio del inglés en los lenguajes informáticos.

Para los desarrolladores, la falta de una interfaz gráfica de usuario (GUI, *Graphical User Interface*) en Hugo es un arma de doble filo. Por un lado, puede significar que los desarrolladores novatos necesiten conocimientos técnicos diferentes de los que ya poseen si su experiencia previa ha sido con plataformas de publicación como WordPress. Por otro lado, también puede significar que los desarrolladores multilingües no estén limitados por una interfaz gráfica de usuario en un solo idioma. Con un generador de sitios estáticos autónomo, el desarrollador puede escapar mejor de los confines técnicos y psicológicos del inglés como idioma predeterminado para la publicación web sin tener que pasar por ningún obstáculo adicional. De lo que no puede escapar, sin embargo, es de que el inglés sea el idioma de programación por defecto. El lenguaje de plantillas sigue estando basado en el inglés y, en el caso de Hugo, la documentación oficial también está disponible únicamente en inglés.

Aunque Hugo tiene suficiente documentación en inglés para la implementación multilingüe, carece de documentación oficial en cualquier otro idioma. El sitio web de Hugo también enlaza a explicaciones y tutoriales más detallados de terceros, incluida una serie de dos partes de Régis Philibert (2018, 2019) sobre cómo hacer un sitio Hugo multilingüe y enlaces a una serie de YouTube sobre Hugo producida por Giraffe

Academy.⁶ Hay otros sitios web similares y canales de YouTube que producen tutoriales de Hugo de alta calidad en otros idiomas, pero no aparecen en ninguna parte del sitio web oficial de Hugo, lo que socava su utilidad como herramienta para interrumpir la supremacía del idioma inglés en la publicación web.

Como ocurre a menudo, los profesionales multilingües llevan la carga de la traducción y la educación, creando tutoriales y traduciendo ellos mismos la documentación del software. Este trabajo es muy valioso y a menudo no se valora lo suficiente. Parte de nuestra motivación para publicar este artículo y guías más detalladas en inglés y español es aliviar la carga de nuestros colegas multilingües y hacer que este trabajo sea más visible para todos en la comunidad. Para ayudar a otros a crear sitios web multilingües con Hugo, compartimos libremente los recursos utilizados para hacerlo en el GitHub del proyecto.⁷ Sarah Tew (2023) también creó tutoriales paso a paso en inglés y español que utilizan una versión simplificada del sitio web *Recetas* como ejemplo y detallan el procesamiento técnico en la siguiente sección. Además, estamos compartiendo este estudio de caso y nuestros hallazgos a través de charlas en conferencias, seminarios web y publicaciones en inglés y español, incluido este artículo.⁸

De hoja de cálculo a sitio web estático

Los sitios web estáticos pueden crearse de muchas maneras a partir de diversos materiales. En esta sección, ofrecemos una visión general de nuestro proceso de transformación de hojas de cálculo en un sitio web para las 53 recetas originales. Se describe cómo fue trabajar en un proyecto bilingüe, se analiza el éxito de este método, se examinan los cambios que deben introducirse en el proceso a medida que aumentamos la ingesta de recetas y se critica el uso de software propietario (Oxygen XML Editor y MS Excel) dentro del proceso (Figura 3).

Recetas se creó a partir de una hoja de cálculo que se transformó, mediante importaciones y transformaciones con XSLT en Oxygen XML Editor, en los archivos Markdown (MD) y parciales HTML que Hugo utilizó para compilar el sitio web final. Utilizamos una hoja de cálculo como base del proyecto porque (1) Melissa Jerome ya había recopilado 53 enlaces a recetas en una hoja de cálculo Excel y (2) las hojas de cálculo son omnipresentes en diversas disciplinas y resultan familiares a personas que, de otro modo,

6. La lista de reproducción de Giraffe Academy está disponible en <https://youtube.com/playlist?list=PLLAZ4kZ9dFpOnyRlyS-liKL5ReHDcj4G3&feature=shared>.

7. Sarah Tew, public-recetas, repositorio GitHub, <https://github.com/SarahTew/public-recetas>.

8. Para obtener una lista de todas las presentaciones del Proyecto de Periódicos Digitales del Caribe y Florida, incluidas las de *Recetas de las Américas*, visite <https://ufndnp.domains.uflib.ufl.edu/presentations/>.

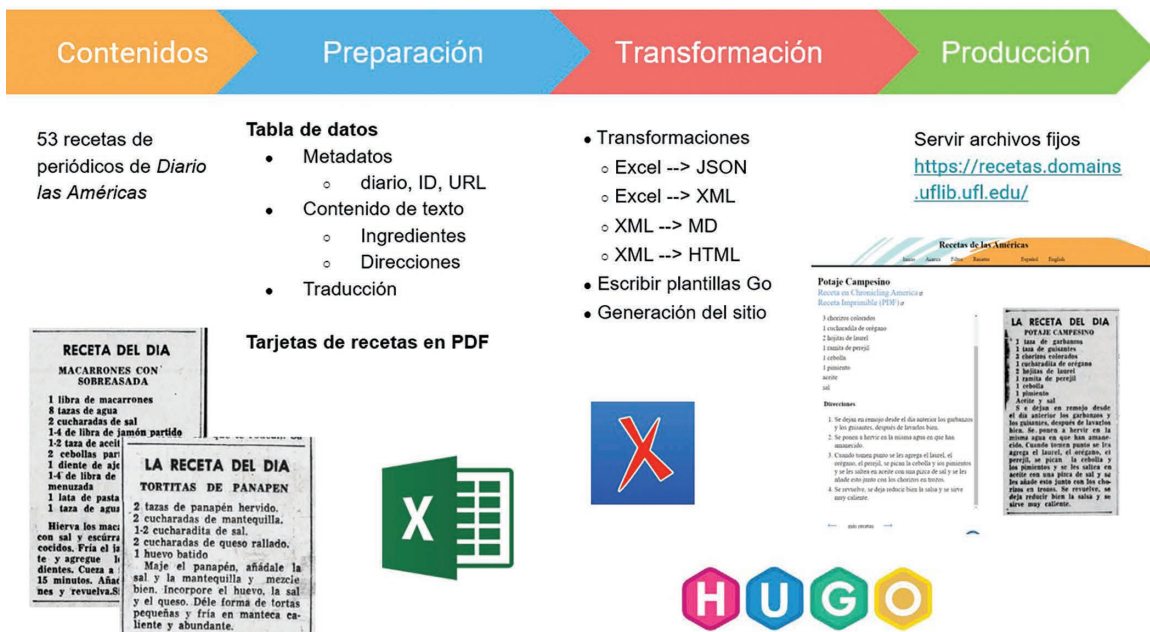


Figura 3. Esquema del proceso para el proyecto *Recetas*.

se sentirían intimidadas por la tecnología digital. Las hojas de cálculo son muy comunes para almacenar información, y desarrollar un método para convertirlas en un sitio web podría abrir posibilidades de publicación web con colegas de las Bibliotecas de la UF y de otros lugares.

Las recetas publicadas actualmente en el sitio web *Recetas* se recopilaban manualmente utilizando la hoja de cálculo original de Melissa Jerome con el título de la receta y un enlace URL a la página del *Diario las Américas* en Chronicling America. Fuimos a cada URL, guardamos un recorte JPEG de la imagen de la receta y copiamos el texto OCR original en español en la hoja de cálculo de Excel. A medida que ampliamos nuestra ingesta de recetas, estamos trabajando en un programa en Python para identificar y recopilar automáticamente las recetas a través de la interfaz de programación de aplicaciones (API, *Application Programming Interface*) de Chronicling America.

Una vez que el texto original OCR en español estuvo en la hoja de cálculo, leímos y limpiamos manualmente el texto. En general, el texto OCR de las recetas estaba en buen estado. Los periódicos en sí se encontraban en su mayoría en un estado entre muy bueno y excelente, con páginas completas libres de manchas, rasgaduras y erratas. El texto OCR fue creado por nuestro proveedor durante la digitalización utilizando ABBYY FineReader Engine. El problema más frecuente en el OCR de las recetas eran las fracciones, que a menudo se representaban como signos de porcentaje (%) en lugar de como dígitos. Los tipos débiles también eran un problema, especialmente con los

caracteres acentuados. Actualmente estamos escribiendo scripts en Python para limpiar, reestructurar y traducir los datos de las recetas.⁹

Tras limpiar y editar el texto OCR en español para que coincidiera con el texto impreso en la imagen del periódico, traducimos el texto al inglés de forma semiautomática. Utilizando la traducción incorporada de Microsoft Excel, primero traducimos al inglés las instrucciones y las celdas de ingredientes, y después hojeamos el texto en inglés y corregimos los errores más importantes. La traducción automática en Excel funcionó mejor de lo que esperábamos, en gran parte debido a la sencillez gramatical y la relativa brevedad del contenido de las recetas del *Diario las Américas*, que consiste casi exclusivamente en listas de nombres concretos seguidas de una serie de frases cortas e imperativas. Dicho esto, nos apoyamos en gran medida en nuestro propio bilingüismo. Los errores en las traducciones de Excel se debían la mayoría de las veces a errores en el texto OCR en español, por lo que al revisar la traducción al inglés, volvíamos a la fuente del error en el texto original en español y lo corregíamos. A medida que ampliamos la ingesta de recetas a otros periódicos, esperamos enfrentarnos a una mayor diversidad lingüística y gramatical, por lo que estamos desarrollando métodos de traducción más robustos en Python. También tenemos curiosidad por investigar el potencial de la traducción como herramienta de limpieza y reestructuración de datos.

El último paso en la preparación del texto dentro de la hoja de cálculo fue extraer manualmente las listas de ingredientes simplificadas y añadir etiquetas para las restricciones dietéticas y el tipo de plato. Esta fue la principal fuente de errores introducidos (por ejemplo, “pan tostado” no se simplificó a “pan”). Actualmente estamos trabajando para semiautomatizar este proceso utilizando Python para limpiar y reestructurar el texto y asignar restricciones dietéticas en función de los ingredientes.

Una vez que el texto de la hoja de cálculo se limpió y estructuró correctamente, se aplicaron hojas de estilo XSLT en Oxygen XML Editor para generar archivos MD y HTML. Los archivos MD (Figura 4) contienen los metadatos de cada receta, incluida una lista simplificada de ingredientes que alimenta la página del filtro, así como instrucciones para que Hugo extraiga los parciales HTML asociados. Los parciales HTML contienen la lista completa de ingredientes y las instrucciones que aparecen en la página web de cada receta (Figura 5). Optamos por utilizar parciales HTML en lugar de incluir el contenido directamente en el archivo MD porque queríamos que los elementos de la página web estuvieran separados y utilizar Hugo para compilarlos. Los archivos de imagen también se almacenan por separado y se llaman mediante un archivo de maquetación HTML. Utilizar parciales HTML y diseños Hugo añadió complejidad

9. Sarah Tew, *recetas2024.ipynb*, Google Colab, 2024, https://colab.research.google.com/drive/1HJ036_cCPkk85XAi6_gKD0-8dIcFAWPt.

```

1 --+
2 id: r39
3 title: gaspacho
4 description: Recetas de las Américas
5 etiquetas: ["vegetariano", "vegano", "sin gluten"]
6 ingredientes: ["ajo", "pimienta en grano", "sal", "aceite", "agua", "vinagre", "tomate", "cebolla", "pepinillos", "pimiento"]
7 plato: ["acompañamiento"]
8 chronam: https://chroniclingamerica.loc.gov/lccn/sn82001257/1955-07-28/ed-1/seq-5/
9 jpg: r39.jpg
10 pdf: r39.pdf
11 ---
12
13 {{< readFile file="./html/recipes/r39.html">}}

```

Figura 4. Archivo Markdown en inglés para la receta 39, Gazpacho.

```

11 <div class="ingredients">
12   <h4 id="ingredhead">Ingredientes</h4>
13   <ul id="ingredients">
14     <li>1 diente de ajo</li>
15     <li>1/4 cucharadita de pimienta en grano</li>
16     <li>1/2 cucharadita de sal</li>
17     <li>3 cucharadas de aceite</li>
18     <li>4 tazas de agua</li>
19     <li>2 cucharadas de vinagre</li>
20     <li>4 tomates picados</li>
21     <li>2 cebollas picadas</li>
22     <li>2 pepinillos picados</li>
23     <li>2 pimientos picados</li>
24   </ul>
25 </div>
26
27
28 <div class="directions">
29   <h4 id="dirhead">Direcciones</h4>
30   <ol id="directions">
31     <li>Muela el ajo junto con la pimienta, la sal y añada el aceite, el agua y el vinagre.</li>
32     <li>Cuele y agregue a las hortalizas picadas.</li>
33     <li>Ponga en la nevera por varias horas y añada cuadritos de pan antes de servirlo.</li>
34   </ol>
35 </div>
36
--

```

Figura 5. HTML parcial en inglés para la Receta 39, Gazpacho.

técnica. Esto fue útil para el proyecto *Recetas*, ya que también queríamos seguir explorando Hugo, practicar la escritura de XSLT, aprender el lenguaje de plantillas, y ser más capaces de hacer sitios web más diversos en el futuro.

Dado que las recetas originales se publicaron en español, el español es el idioma por defecto del sitio web de *Recetas* publicadas. Los campos de metadatos en spanglish ilustran algunos hábitos de una persona multilingüe que trabaja en dos idiomas. Sarah Tew escribió el XSLT para transformar la hoja de cálculo en los archivos MD y HTML utilizados por Hugo para generar el sitio. Los archivos MD contienen metadatos sobre cada receta, incluido el título, el periódico y la lista de ingredientes. El sitio web en sí se desarrolló de forma bilingüe, con Sarah cambiando entre español e inglés dependiendo de la fase del proyecto y del material con el que estuviera trabajando. El contenido de la receta en sí estaba en español, pero el editor XML Oxygen, la documentación Hugo y

el lenguaje de plantillas Go estaban todos en inglés, lo que la llevó de un lado a otro del español y del inglés a lo largo del desarrollo. La naturaleza bilingüe del proyecto se evidencia en la mezcla de idiomas dentro de los archivos MD, con algunos campos, como “Ingredientes” y “Etiquetas”, en español y otros, como “Title” o “Paper”, en inglés. Cambiaremos todos los campos al español con la próxima ingesta y actualización a gran escala para que el proyecto sea más utilizable y replicable para otros. Los generadores de sitios estáticos admiten fácilmente metadatos en varios idiomas. Aunque pensamos que, en última instancia, esto no es deseable para *Recetas*, puede ser útil para otros proyectos que se beneficiarían de la flexibilidad de nombrar y mezclar campos en varios idiomas.

Una vez creados los archivos MD, los parciales HTML y escritos los borradores de las plantillas para la maquetación, revisamos el sitio a través del comando de *local host* incorporado en Hugo. El sitio alojado localmente fue una herramienta cómoda para revisar las traducciones y detectar errores en las listas de ingredientes y etiquetas tanto en español como en inglés. La página de filtros, que enumera los ingredientes simplificados por orden alfabético, resultó útil para encontrar erratas e ingredientes simplificados incorrectamente (por ejemplo, “sal” por “sal gruesa”). Además de ser útiles para encontrar erratas, la flexibilidad de los sitios estáticos para cambiar ágilmente la forma en que se utiliza y se muestra el contenido los convierte en herramientas valiosas para la gestión y exploración de contenidos, incluso para proyectos que no tienen intención de publicar nunca un sitio web. Las plantillas de cada página pueden editarse rápidamente para mostrar información diferente; información en formatos distintos, como listas y párrafos; e información ordenada de forma diferente, como por recuento o por fecha.

Una vez satisfechas con el borrador del sitio web, generamos todos los archivos necesarios para el sitio web desde la línea de comandos con Hugo en poco más de un segundo. A continuación, subimos estos archivos a nuestro servidor institucional, Lib-Domains, utilizando Core FTP y alojamos el sitio web.¹⁰

Software propietario

Debido a los recursos institucionales para el espacio del servidor, el alojamiento, las licencias de software y los ordenadores, el costo total para el proyecto subvencionado fue de 0 \$. Aunque logramos nuestros objetivos con un costo financiero directo mínimo, sólo tuvimos un éxito parcial a la hora de alinear *Recetas* con los principios de asequibilidad y uso de herramientas de código abierto. En cuanto al software utilizado para *Recetas*, Hugo y Core FTP LE son gratuitos y de código abierto, pero Microsoft Excel

10. University of Florida George A. Smathers Libraries, UFLib Domains, consultado el 14 de abril de 2024, <https://domains.uffib.ufl.edu>.

y Oxygen XML Editor están patentados y hay que comprarlos. El uso de herramientas de código abierto va mucho más allá en el apoyo a la igualdad de acceso a todos los profesionales de las HD. A medida que ampliamos la ingesta de recetas, estamos eliminando el formato propietario de Excel del proceso en favor de los archivos de valores separados por comas (CSV). En su lugar, tenemos previsto utilizar Python para limpiar, estructurar y traducir el texto OCR en bruto.

Al contrario que con Excel, no tenemos planes de eliminar Oxygen XML Editor de nuestro proceso. En este momento, ya estamos familiarizadas con Oxygen XML Editor, y su uso nos libera tiempo y espacio mental para ampliar la colección, desarrollar las funciones del sitio web y explorar los datos de las recetas, todas ellas prioridades más importantes en este momento. Consideramos la posibilidad de cambiar a alternativas más baratas o de código abierto como EditiX y Treeline, pero no nos planteamos seriamente estas alternativas ya que ninguna de ellas era una sustitución lo suficientemente sencilla y no apoyaba significativamente el crecimiento de *Recetas*.¹¹ En este caso, los objetivos y recursos de nuestro proyecto específico pesaron más que nuestro deseo de utilizar software de código abierto.

Oxygen XML Editor también acorta la distancia entre el generador de sitios estáticos y el desarrollador de forma eficiente, suponiendo que conozcan uno de los lenguajes soportados por la GUI. Oxygen XML Editor es una herramienta útil para gestionar directorios y ensamblar los archivos del proyecto de una forma más familiar para desarrolladores inexpertos. Hugo también puede escribir nuevos archivos, listar directorios y mover y copiar archivos, pero se basa en el uso de la interfaz de línea de comandos, que a menudo intimida a los creadores novatos. En este sentido, la interfaz gráfica de Oxygen es una ventaja. La navegación por directorios tiene un aspecto y un funcionamiento similares a los del Explorador de archivos de Windows, las herramientas están disponibles a través de menús desplegables estándar y una cinta de iconos ayuda a los usuarios a encontrar rápidamente las funciones principales.

Los escenarios de transformación de Oxygen fueron esenciales para *Recetas*. Creamos varias hojas de estilo XSLT y escenarios para generar los archivos MD y los parciales HTML en español e inglés. Estos se encuentran disponibles en el repositorio GitHub.¹² Oxygen incluye un depurador XSLT, y los escenarios de transformación fueron más fáciles de usar y entender que transformar en otro programa.

Sin embargo, la interfaz gráfica de usuario no es igual de accesible para todos los usuarios. Los idiomas disponibles de Oxygen se limitan al inglés (por defecto), alemán,

11. EditiX (<https://www.editix.com>) es de código abierto pero no gratuito, aunque es mucho más barato que Oxygen: una única compra de 19 dólares por los binarios de licencia académica frente a la suscripción anual de 68 dólares de Oxygen. Treeline (<https://treeline.bellz.org>) es gratuito y de código abierto.

12. Sarah Tew, public-recetas, repositorio GitHub, <https://github.com/SarahTew/public-recetas>.

francés, japonés y holandés. Es posible añadir el español y otros idiomas, pero para ello es necesario crear un archivo de localización de la interfaz (Oxygen XML Editor, s.f.). Es una pena que esta carga recaiga sobre los usuarios y resulta sorprendente que un idioma mundial como el español no esté soportado automáticamente en el paquete básico de Oxygen. No hay que descartar la barrera adicional que supone utilizar software en otro idioma o el tiempo y la energía necesarios para adaptarse a otro idioma. Esperamos que Oxygen añada el español como idioma incorporado en el futuro. Hasta entonces, su valor es menor para los profesionales de idiomas no compatibles.

Conclusión

Es vital invertir tiempo y energía en herramientas que puedan ayudar a que la publicación electrónica sea más accesible, especialmente para quienes publican en idiomas distintos del inglés y fuera de las instituciones tradicionales o bien dotadas de recursos en el Norte Global. Los sitios estáticos son una opción viable para la publicación web de bajo costo para proyectos de HD según los principios de computación mínima. A diferencia de otros generadores de sitios estáticos populares, el manejo de idiomas incorporado en Hugo disminuye las dependencias y la complejidad de los sitios web multilingües. Esto ayuda a reducir las barreras a la publicación académica multilingüe, aumentando potencialmente la participación de comunidades históricamente marginadas como editores y lectores.

El uso de Hugo, Excel y Oxygen XML Editor para crear el proyecto web bilingüe *Recetas de las Américas* fue en gran medida un éxito, pero hay que seguir trabajando para reducir la dependencia del software propietario y maximizar el potencial de reutilización, especialmente para los profesionales de HD con pocos recursos. Actualmente se está trabajando para eliminar Excel del proceso, pero sustituir las funciones XSLT de Oxygen XML Editor y aprender un nuevo software implica un alto gasto de tiempo, sobre todo teniendo en cuenta el apoyo institucional a *Recetas* de las Bibliotecas de la Universidad de Florida.

La creación de un sitio web estático sencillo con Hugo sólo requiere conocimientos básicos de tecnologías web, pero la utilización completa del generador para crear un proyecto web más complejo requiere conocimientos intermedios. La parte más difícil de usar Hugo fue el lenguaje de plantillas; sin embargo, la flexibilidad de las plantillas junto con el soporte de idiomas de Hugo puede significar que los sitios web multilingües sean más fáciles de hacer que con otros métodos. La falta de documentación de Hugo en idiomas distintos del inglés hace que el software sea menos accesible. A medida que continuemos expandiendo *Recetas*, seguiremos compartiendo libre y abiertamente nuestros métodos, resultados y tutoriales en español e inglés.

Referencias

- Díaz, Cris. 2018. "Using Static Site Generators for Scholarly Publications and Open Educational Resources". *Code4Lib*, no. 42. <https://journal.code4lib.org/articles/13861>.
- Dombrowski, Quinn. 2022. "Minimizing Computing Maximizes Labor". *Digital Humanities Quarterly* 16, no. 2. <http://www.digitalhumanities.org/dhq/vol/16/2/000594/000594.html>.
- Gil, Alex, y Élika Ortega. 2016. "Global Outlooks in Digital Humanities: Multilingual Practices and Minimal Computing". En *Haciendo humanidades digitales: Practice, Training, Research*, editado por Constance Crompton, Richard Lane y Ray Siemens. Londres: Routledge. <https://doi.org/10.4324/9781315707860>.
- Hugo. s.f.-a. "Hugo Themes". Consultado el 15 de abril de 2024. <https://themes.gohugo.io>.
- . s.f.-b. "Multilingual Mode". Consultado el 15 de abril de 2024. <https://gohugo.io/content-management/multilingual/>.
- Jerome, Melissa. 2022. "Recetas de las Américas/Recipes from the Americas". US Caribbean & Florida Digital Newspaper Project, 5 de octubre de 2022. <https://ufndnp.domains.uflib.ufl.edu/recetas/>.
- Library of Congress. s.f. "Chronicling America". Consultado el 31 de diciembre de 2023. <https://chroniclingamerica.loc.gov>.
- Lincoln, Matthew, Jennifer Isasi, Sarah Melton y François Dominic Laramée. 2022. "Relocating Complexity: The Programming Historian and Multilingual Static Site Generation". *Digital Humanities Quarterly* 16, n° 2. <http://www.digitalhumanities.org/dhq/vol/16/2/000585/000585.html>.
- NDNP (Programa nacional de periódicos digitales). 2019. "About the Program". Library of Congress. Última modificación: 14 de marzo de 2019. <https://www.loc.gov/ndnp/about.html>.
- . 2023. "Technical Guidelines for Applicants". Library of Congress. Última modificación: 25 de julio de 2023. https://www.loc.gov/ndnp/guidelines/NDNP_202426TechNotes.pdf.
- Newson, Kaitlin. 2017. "Tools and Workflows for Collaborating on Static Website Projects". *Code4Lib*, no. 38. <https://journal.code4lib.org/articles/12779>.
- Nilsson-Fernández, Pedro, y Quinn Dombrowski. 2022. "Multilingual Digital Humanities". En *The Bloomsbury Handbook to the Digital Humanities*, editado por James O'Sullivan, 83–92. Londres: Bloomsbury. <https://www.bloomsbury.com/us/bloomsbury-handbook-to-the-digital-humanities-9781350232112/>.
- Oxygen XML Editor. s.f. "Localizing the User Interface". Consultado el 31 de diciembre de 2023. <https://www.oxygenxml.com/doc/versions/26.0/ug-editor/topics/changing-user-interface-language.html>.
- Philibert, Régis. 2018. "Hugo Multilingual Part 2: Strings Localization". 23 de agosto de 2018. <https://www.regisphilibert.com/blog/2018/08/hugo-multilingual-part-2-i18n-string-localization/>.
- . 2019. "Hugo Multilingual Part 1: Content Translation". Última modificación 16 de febrero de 2019. <https://www.regisphilibert.com/blog/2018/08/hugo-multilingual-part-1-managing-content-translation/>.
- Risam, Roopika. 2018. *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Chicago: Northwestern University Press.
- Risam, Roopika, y Alex Gil. 2022. "Introduction: The Questions of Minimal Computing". *Digital Humanities Quarterly* 16, n° 2. <https://www.digitalhumanities.org/dhq/vol/16/2/000646/000646.html>.
- Simple IT Rocks. 2019. "Multilingual Jekyll without Plugins". Última modificación 27 de marzo de 2019. <https://simpleit.rocks/ruby/jekyll/tutorials/multilingual-jekyll-without-plugins/>.
- Tanasescu, Raluca. 2021. "Complexity and the Place of Translation in Digital Humanities". En *Exploring the Implications of Complexity Thinking for Translation Studies: Post-disciplinary Communities of Practice in the Translation Studies Network*, editado por Kobus Marais y Reine Meylaerts, 30–72. Londres: Routledge. <https://doi.org/10.4324/9781003105114-3>.

- Tew, Sarah. 2023. “De una hoja de cálculo a un sitio web multilingüe utilizando el editor XML de Oxygen y el generador de sitios estáticos Hugo”. <https://sites.google.com/ufl.edu/dlf-2023-es/home>. En inglés como “From Spreadsheet to Multilingual Website Using Oxygen XML Editor & Hugo Static Site Generator”. <https://sites.google.com/ufl.edu/dlf2023/home>.
- . s.f. public-recetas. Repositorio GitHub. <https://github.com/SarahTew/public-recetas>.
- Tew, Sarah y Melissa Jerome. s.f. *Recetas de las Américas*. <https://recetas.domains.uflib.ufl.edu>.
- USCFDNP (US Caribbean & Florida Digital Newspaper Project). s.f. “About the Project”. Consultado el 17 de junio de 2024. <https://ufndnp.domains.uflib.ufl.edu/about/>.
- Varner, Stewart. “Minimal Computing in Libraries: Introduction”. *Minimal Computing*, 15 de enero de 2017. <https://go-dh.github.io/mincomp/thoughts/2017/01/15/mincomp-libraries-intro/>.