



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
POSGRADO EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN  
FACULTAD DE FILOSOFÍA Y LETRAS  
INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y DE LA INFORMACIÓN

MODELO DE INTEGRACIÓN DE METADATOS PARA PRESERVACIÓN EN LA EDICIÓN DEL  
LIBRO DIGITAL

**TESIS**  
QUE PARA OPTAR POR EL GRADO DE:  
DOCTORA EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN

PRESENTA:  
**ANA YURI RAMÍREZ MOLINA**

TUTOR O TUTORES PRINCIPALES  
DR. JUAN VOUTSSÁS MÁRQUEZ, INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y  
DE LA INFORMACIÓN

MIEMBROS DEL COMITÉ TUTOR  
DRA. BRENDA CABRAL VARGAS, INSTITUTO DE INVESTIGACIONES BIBLIOTECOLÓGICAS Y  
DE LA INFORMACIÓN  
DRA. ISABEL GALINA RUSSELL, INSTITUTO DE INVESTIGACIONES BIBLIOGRÁFICAS

Ciudad Universitaria, Ciudad de México, febrero del 2022



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

**Agradecimientos.**

A mis padres por su eterna confianza y su incansable apoyo.  
A mi hijo por ser mi inspiración para dedicarme a este sueño.  
A Miguel Angel por el tiempo que pudo ser un compañero y siempre un amigo.

A mis maestros, la Dra. Hanna Oktaba por sus enseñanzas, el Dr. Juan Voutssás por su paciencia, a las Doctoras Isabel Galina, Brenda Cabral y Martha Romero por guiarme. Gracias a todos, porque sus enseñanzas me permitieron encontrar el camino para hacer un trabajo como este.

A la UNAM por cobijarme como estudiante y como trabajadora, dirigiéndome a este nuevo campo de estudio que tanto me apasiona.

Y a Dios, porque la vida me ha alcanzado para cumplir con este sueño a pesar de tantas adversidades.

Dedico este trabajo a mi hijo Angel Cipatli, porque la felicidad, satisfacción y orgullo que siento al ser su mamá, es lo que me mantuvo firme y me ayudó a no rendirme en el camino y por eso, logré cumplir un sueño más, en el cual, él siempre estuvo a mi lado dándome la fuerza que necesitaba, sólo con ser mi hijo.

Es poco lo que he alcanzado en esta vida académica comparado con las satisfacciones que he tenido en mi vida como madre, hija y confidente, sin embargo, este logro representa momentos en mi vida de una ardua lucha, por eso, me hace feliz compartir el triunfo contigo.



**CONTENIDO**

<b>Introducción</b> .....	5
<b>I. Metodología.</b> .....	9
<b>Capítulo I. El libro digital y su preservación.</b> .....	12
<b>1.1. El objeto libro.</b> .....	12
<b>1.2. Origen del libro digital</b> .....	16
1.2.1. Visión general.....	16
1.2.2. Cronología del avance tecnológico que impactó en el desarrollo del libro digital .....	18
1.2.3. Cronología del desarrollo del libro digital .....	23
1.2.4. Análisis del origen del libro digital .....	30
<b>1.3. Esencia del libro digital como software</b> .....	31
1.4. Legislación para la preservación del libro digital .....	39
<b>1.5. Origen de la preservación digital</b> .....	48
<b>1.6. La preservación digital y el libro digital</b> .....	53
1.6.1. Organizaciones y trabajos sobresalientes en preservación digital .....	53
1.6.2. Modelo OAIS y otras perspectivas.....	57
1.6.3. La preservación del libro digital .....	62
<b>1.7. Metadatos para preservación y de preservación.</b> .....	65
1.7.1. El concepto de metadato .....	66
1.7.2. Metadatos de preservación .....	72
1.7.3. Metadatos para preservación .....	75
1.7.4. La problemática de los metadatos para preservación .....	79
<b>Capítulo II. La edición del libro digital.</b> .....	83
<b>2.1. La edición, ciclo de vida y cadena de valor del libro</b> .....	83
<b>2.2. La esencia de la Ingeniería de Software aplicable al libro.</b> .....	91
<b>2.3. La edición de libros digitales y la selección de casos para su análisis</b> .....	99
<b>2.3.1. Guía de libros electrónicos CEPAL (caso 1).</b> .....	103
<b>2.3.2. Edición digital de textos literarios (caso 2).</b> .....	106
<b>2.3.3. Introducción a la edición digital (caso 3).</b> .....	109
<b>2.3.4. Edición de publicaciones Web (caso 4).</b> .....	110
<b>2.4. Análisis de casos para la integración de un proceso de edición.</b> .....	113
<b>2.5. Aplicación de la esencia de Ingeniería de Software en la edición del libro digital.</b> .....	118

2.5.1. Etapa de: Control y Seguimiento. ....	119
2.5.2. Etapa de: edición.....	125
2.5.3. Etapa de: Programación.....	128
2.5.4 Etapa de: Distribución.....	134
2.5.5 La especificación de un proceso de edición fortalecido por la esencia de la Ingeniería de Software. ....	138
<b>Capítulo III. Desarrollo del Modelo.</b> .....	141
3.1. ¿Qué es un modelo? .....	141
3.2. Bases del modelo descrito en este trabajo de investigación.....	145
3.3. Especificación frente a implementación.....	146
3.4. Abstracción frente a detalle.....	152
3.5. Recomendaciones de aplicación e implementación del modelo.....	162
3.5.1. De la aplicación como editor.....	163
3.5.2. De la aplicación en entidades que resguardan libros digitales.....	164
<b>Conclusiones.</b> .....	166
<b>Referencias.</b> .....	171
<b>Anexos</b> .....	178
I. Índice de imágenes e ilustraciones .....	178
II. Índice de tablas .....	179
III. Diagramas del Proceso Editorial elaborado.....	180
IV. El modelo .....	181

## INTRODUCCIÓN.

De acuerdo con las directrices para la salvaguarda del patrimonio documental de la UNESCO, el libro digital o electrónico forma parte del patrimonio documental por lo que es responsabilidad de las instituciones depositarias como bibliotecas o archivos nacionales, su resguardo y cuidado (UNESCO, 2003, p. 6). También es considerado patrimonio digital, de ahí que las *Directrices para la preservación digital*, recalcan la importancia de apoyar a las instituciones depositarias en el combate de los factores que lo hacen vulnerable por su naturaleza intangible (UNESCO & BNA, 2003, p. 3).

Por lo anterior, dichas directrices promueven una “política nacional de preservación” (UNESCO & BNA, 2003, p. 13) en los países, que incluye desde la modificación o adecuación de las legislaciones como el Depósito Legal (IPA, 2014) hasta el establecimiento de modelos, criterios y procesos técnicos y administrativos dentro de las instituciones depositarias (bibliotecas y archivos) para orientar a los editores y/o creadores de publicaciones digitales y así generar garantías para la protección de los derechos del editor y mecanismos efectivos para la recepción de las publicaciones con mejores condiciones para la preservación digital.

También, se sabe que para establecer procesos de preservación digital es necesario adaptarlos al medio social en el que los materiales o libros son creados, adecuando dichos procesos a las necesidades de cada nación o institución y a las posibilidades económicas, geográficas, jurídicas, políticas y técnicas de las mismas (Voutssás, 2011). Y existe un fuerte impulso o tendencia en el ámbito de la preservación digital para aprovechar la mayor cantidad de información producida con una publicación y así, reducir la cantidad de recursos humanos empleados para la identificación o generación de datos dentro de las instituciones depositarias (Owens, 2018).

Dicha tendencia se conoce como *More Product, Less Process* o MPLP (Owens, 2018) y promueve el control y organización de los materiales digitales con una menor inversión de recursos, aprovechando al máximo la información producida con la publicación digital, ya que no hacerlo, genera grandes tareas de investigación para recabar información acerca de los materiales digitales y en muchos casos, a pesar del gran esfuerzo es imposible obtener lo requerido para el resguardo de los mismos.

En concordancia, se han desarrollado modelos para preservación digital, a nivel mundial, que resaltan la importancia de los metadatos (descriptivos, documentales, de contexto, estructurales, funcionales, jurídicos, administrativos, tecnológicos, de ubicación, de autenticidad, de preservación, de privilegios, de procedimiento, etc.) pues facilitan la organización y control dentro de las entidades depositarias o encargadas del resguardo de documentos, libros y/o publicaciones digitales.

En contraste con esto, en 2018, la Biblioteca Jurídica del Congreso de los Estados Unidos, publicó un informe en el que analiza a 15 países del mundo, seleccionados por tener una amplia experiencia con la recolección del Depósito legal digital y por lo tanto en la preservación de publicaciones y libros digitales, del cual se puede resaltar lo siguiente (LLC-EUA, 2018):

- En la mayoría de los países encuestados la responsabilidad del depósito recae en los editores o productores y se establecen normas precisas para la entrega de estos materiales en repositorios y con metadatos.
- La mayoría de los países encuestados extienden la protección a sitios web, publicaciones en línea que incluye a las publicaciones o libros nacidos digitales y copias digitalizadas de materiales impresos.

De ahí que, la Biblioteca del Congreso de los Estado Unidos en dicho informe resalta entre otras cosas, la participación de los editores o productores de publicaciones digitales con las entidades depositarias para la integración de políticas, acuerdos, modelos y legislaciones que ayuden dentro de sus naciones al resguardo y protección del patrimonio digital.

Por otro lado, en México, en el año 2020 la recepción del Depósito Legal Digital se hizo obligatorio derivado de la promulgación de una *Ley General de Bibliotecas* y esto preocupó a los editores, quienes se han mostrado inconformes y algunos incluso, han pensado en recurrir al amparo para evitar sanciones (Letralia, 2021), de ahí que las instituciones depositarias siguen en la búsqueda de acuerdos. Cabe resaltar que dicha ley no menciona información referente a la recepción de metadatos, sin embargo, en el sitio web de la Biblioteca Nacional de México, existen formatos con la solicitud de estos para la recepción del Depósito Legal (BNM, 2021).

En contraste, también en México los editores reconocen que la cadena de valor por la que transcurre el libro ha cambiado por la incorporación de nuevas tecnologías que incluyen la fabricación de software, medios electromagnéticos y virtuales (Piedras, 2013, p. 35), de ahí que ahora, la industria editorial para la producción de libros electrónicos se ve relacionada íntimamente con el desarrollo de software y en consecuencia con la ingeniería de software, pues como se verá en esta investigación, es el área de conocimiento encargada de estudiar y proponer los mecanismos y prácticas adecuadas para la producción y mantenimiento de productos de software e incluye al libro digital.

Además, de acuerdo con un estudio elaborado en el año 2020, acerca del sector editorial en México, es necesario integrar y sistematizar las transacciones entre los distintos participantes de la cadena de valor del libro, así como definir protocolos y buenas prácticas de intercambio de información e implementar sistemas de gestión e intercambio de metadatos (G&G, 2020, p. 87) ya que esto facilita la comercialización y difusión de las publicaciones, pero además como se verá en este trabajo de investigación, son indispensables para la preservación a largo plazo, pues cabe resaltar que la industria del

libro es “el undécimo mercado más grande del mundo para el sector editorial” (G&G, 2020, p. 89) de ahí su importancia.

A pesar de todo lo anterior, aún no existen modelos para la integración de metadatos para preservación de libros digitales y los editores están conscientes de que requieren mucho apoyo para reducir la brecha digital pues la capacidad tecnológica instalada para el trabajo editorial está concentrada sólo en algunos estratos de producción (G&G, 2020, p. 15) de ahí que los editores necesitan de profesionales especializados en el desarrollo de software para la creación, distribución y comercialización de libros y no siempre es posible costear esa inversión, por lo que en este trabajo de investigación entre otras cosas, se resaltarán las ventajas del uso de buenas prácticas de Ingeniería de software en el ámbito editorial para la construcción de libros digitales.

Por lo tanto, la investigación que aquí se presenta, recalca la importancia que tienen los metadatos para la preservación del libro digital e identifica cuáles son básicos y necesarios, por lo que es conveniente propiciar su integración durante la edición, además de proponer el uso de herramientas y buenas prácticas de ingeniería de software para la creación de un modelo que orienta a los productores de libros en la integración de metadatos durante el proceso de edición, el cual podría emplearse principalmente en México ya que como se mencionó, el sector editorial reconoce la falta de buenas prácticas para el intercambio de metadatos.

De todo lo anterior, por tanto, se tiene identificado lo siguiente:

- La creación de libros digitales con metadatos aún no es una práctica común en la industria editorial mexicana, contrario a lo que sucede en países que llevan a cabo preservación digital en el mundo.
- La preservación digital requiere un registro vasto de datos.
- No existen modelos de integración de metadatos para preservación durante la edición de los libros digitales en México.
- Está comprobado que la preservación de los libros digitales requiere la colaboración de los productores y/o creadores, con las instituciones receptoras del Depósito Legal o depositarias.
- Las buenas prácticas de Ingeniería de Software son efectivas para la construcción de productos de software, por lo que su uso en la producción de libros digitales puede ser de utilidad.

Así pues, este trabajo busca resolver las siguientes preguntas de investigación:

1. ¿Cuáles son los metadatos básicos que deben generarse durante la edición del libro digital, para propiciar su preservación?
2. ¿Cuál es el papel que juegan los editores y entidades receptoras de depósito legal en la especificación de metadatos para preservación digital?

3. ¿La adopción de buenas prácticas de ingeniería de software en la edición del libro digital, permitiría generar los metadatos necesarios para su preservación?

Por todo lo anterior, esta investigación tiene por objetivo general, crear un modelo de integración de metadatos para preservación en la edición del libro digital a partir de los siguientes objetivos específicos:

1. Definir cuáles son los metadatos básicos que se deben generar durante la edición para propiciar la preservación del libro digital.
2. Identificar qué papel juegan los editores y entidades receptoras de depósito legal en la especificación de metadatos para preservación digital.
3. Determinar cómo la adopción de buenas prácticas de ingeniería de software en la edición del libro digital permite generar los metadatos necesarios para preservación.

De todo esto se deriva las siguientes hipótesis de investigación:

1. Es posible establecer los metadatos básicos para preservación de un libro digital y generarlos durante su edición con la participación de los editores y entidades de Depósito Legal.
2. Es posible construir un modelo que considere la integración de metadatos para preservación durante la edición del libro digital a partir de la aplicación de buenas prácticas de ingeniería de software.

Este trabajo se realizó aplicando un método de investigación cualitativo basado en un diseño de teoría fundamentada, pues se produjo un modelo que explica de forma general la importancia de la integración de metadatos para preservación en la edición del libro digital a partir del estudio de casos tipo y por oportunidad (Hernández S., 2014, p. 472) aplicando buenas prácticas de ingeniería de software, del cual se habla con detalle en el siguiente capítulo.

## I. METODOLOGÍA.

Para esta investigación el método cualitativo se aplicó en tres etapas, la primera concentró desde el planteamiento del problema hasta la identificación o definición de conceptos clave para este trabajo, la segunda etapa fue para la especificación del método de recolección de datos y el análisis mismo de lo que se identificó con la revisión de la literatura en la etapa anterior y la última etapa se concentró en el desarrollo del modelo generado como un producto de esta investigación y derivado de las etapas anteriores, ver Figura 1.

Figura 1. Método Cualitativo.



Fuente: La autora, octubre 2021.

De ahí que, la primera etapa partió del planteamiento del problema en donde se hizo el análisis de la importancia de los metadatos y una revisión holística de la literatura referente al libro digital y su esencia como libro y como software, con el propósito de construir una aproximación al contexto en el que se desarrolla su preservación y resaltar el porqué de la importancia de los metadatos.

También fue necesario revisar la legislación conocida como Depósito Legal, principalmente en los países en los que la preservación digital de libros tiene un avance significativo y además se revisó el estado del Depósito Legal en México, para poder comparar cuáles eran las implicaciones o diferencias más importantes, con lo cual, se confirmó la importancia de los metadatos y la relevancia de la participación de los editores.

En esta etapa se contempló el análisis de la edición del libro digital a partir del estudio de su cadena de valor, para entender cómo ha cambiado la edición derivada de las características que adquiere el libro digital, por su esencia como software y con lo cual se genera una relación en su edición o construcción, con la ingeniería de software.

En la segunda etapa se estableció, para este estudio, la utilización de “una muestra de casos tipo y por oportunidad” (Hernández S., 2014, pp. 387-389), pues se trató de casos que oportunamente fueron identificados durante la revisión de la literatura, por aportar riqueza en la variedad de perspectivas de edición, ya que describen cuatro fenómenos diferentes en la edición de libros o publicaciones digitales:

- Un editor de materiales impresos que integra la creación de libros digitales en su proceso.
- La experiencia de autoedición desde la perspectiva de un autor
- La edición de un libro digital
- Las recomendaciones de una entidad emisoras de estándares para publicaciones digitales.

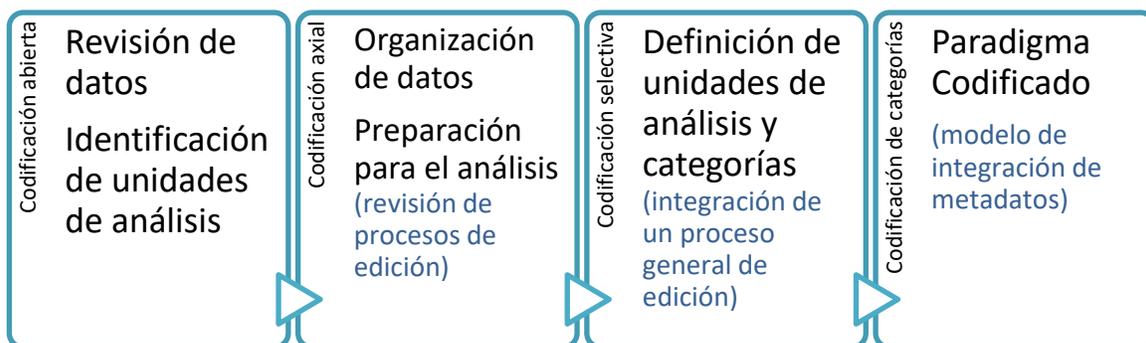
Además, para la recolección de datos se consideraron las siguientes “unidades de análisis” (Hernández S., 2014, p. 397):

- **Significados:** para entender los conceptos y entidades relacionadas en la cadena de valor del libro y en consecuencia en la edición.
- **Prácticas:** para analizar las actividades conductuales involucradas en cada caso de edición de libro o publicación seleccionado.
- **Papeles y roles:** para entender las unidades de organización que intervienen con un sentido y significado, generando vínculos y tipologías en un ámbito organizacional en el que se da la edición de una publicación o libro digital.
- **Organizaciones y procesos:** para entender cuáles son las tareas que se dan de forma sucesiva y con el fin de producir una publicación o libro digital.

Para el análisis detallado de los datos se ocupó un procedimiento señalado por Hernández Sampieri como “teoría fundamentada” (Hernández S., 2014, p. 422), pues a partir de utilizar las prácticas de ingeniería de software, los datos encontrados se fueron organizando en un proceso iterativo, conocido en la investigación cualitativa como “diseño sistemático”, con el cual el estudio de la información recolectada se divide en cuatro procesos “codificación abierta”, “codificación axial”, “codificación selectiva”, “codificación de categorías” con lo cual es posible generar el “paradigma codificado” (Hernández S., 2014, p. 474) que en este caso, es un modelo para la integración de metadatos.

De lo anterior que el análisis para este trabajo de investigación comenzó desde la definición de las unidades de información que se fueron identificando en cada ejemplo de edición de libro o publicación seleccionado, para después derivar en la construcción de un modelo que refleja cómo se da la integración de metadatos para preservación durante la edición del libro digital, ver Figura 2.

Figura 2. Diseño sistemático para el análisis



Fuente: La autora, octubre 2021.

Cabe señalar que, en esta segunda etapa, además de llevar a cabo un proceso de análisis, el uso de herramientas de codificación empleadas para la abstracción de los datos identificados, partió de buenas prácticas empleadas en la Ingeniería de Software y así, obtener un diagrama del proceso de edición integrado de las cuatro perspectivas seleccionadas, que sirvió en la tercera etapa como base para la codificación el modelo de integración de metadatos para preservación en la edición del libro digital.

La tercer y última etapa contempló la construcción del modelo a partir de los resultados obtenidos en la primera y segunda etapa, con la aplicación de prácticas de ingeniería de software y basado en la construcción de un modelo que parte de la aplicación de conceptos establecidos en FRBR, el modelo OAIS y criterios de calidad de software, en el cual se establece la participación de los editores y entidades receptoras de depósito legal en la definición de metadatos para preservación y además, se especifican cuáles son los metadatos básicos para preservación desde la perspectiva del libro digital como software.

## CAPÍTULO I. EL LIBRO DIGITAL Y SU PRESERVACIÓN.

“Definir el libro electrónico es un trabajo aún más arduo al hallarnos en el filo de la navaja tecnológica de lo que va a ser, pero que aún no es, definiendo ese «es» como algo que aún no existe en su forma definitiva, pero que existirá en los próximos años”  
(Toledo Morales & Sánchez, 2002, p. 127)

En este primer capítulo se hace una revisión del libro digital desde su esencia como objeto libro, además se explora su origen en el ámbito tecnológico con un recorrido histórico para entender cuáles fueron los hitos que contribuyeron o fueron específicamente, orientados a su nacimiento. También se explora su naturaleza como software y su definición a partir de las características que se derivan de dicha esencia intangible.

Además, se explican los conceptos relacionados con su preservación y el impacto de esto en el depósito legal y cómo funciona en países que están haciendo preservación digital, comparado con lo hecho en México, para entender qué hace falta, con lo cual, también se resalta la importancia de los metadatos, para concluir abordando estos con detalle y así establecer las tipologías o clasificaciones que son necesarias para lograr la preservación de los libros digitales.

### 1.1. EL OBJETO LIBRO.

La palabra libro, en su etimología se refiere esencialmente al medio material con el que se constituyó en sus orígenes, como medio manuscrito, ya que ésta, se compone de dos vocablos que describen a la corteza y lugar en el que se obtenía el papel o papiro (Cruz Quintana, 2015).

Vocablo latín liber [...] entretela que se halla entre el tronco y la corteza, con lo que se hacía lo que ahora llamamos papel [...] Biblos [...] ciudad libanesa donde se encuentra el papiro, que es una planta acuática de la que se obtiene un material homónimo, donde posteriormente se puede escribir (Cruz Quintana, 2015, p. 305).

De ahí que, el análisis del objeto libro para entender al libro digital, no puede partir del simple hecho de entender el origen de su nombre, pues de entrada el libro digital no se compone de papel. Sin embargo, como veremos más adelante, el objeto libro no se define sólo por su composición material de ahí que el uso de la palabra libro, va más allá de su origen etimológico.

Cabe resaltar que durante el desarrollo de este trabajo se identificaron definiciones empleadas cotidianamente para el libro y que de alguna manera también dejan fuera al libro digital, pues se concentran principalmente en las características físicas de los impresos, por ejemplo, la definición de la UNESCO, que lo refiere como “toda publicación impresa no

periódica de al menos 49 páginas excluyendo cubiertas” (UNESCO, citado por Voutssás, 2005, p. 62).

Esta definición, incluso es catalogada como arbitraria, por especialistas como Fernando Cruz, pues desde su perspectiva, el número de páginas no encierra ningún sentido claro respecto al libro (Cruz Quintana, 2015, p. 305), lo cual también aplica para el libro digital, pues como se verá más adelante incluso el concepto de páginas cambia con la virtualidad, sin embargo, lo rescatable de esta definición, desde el punto de vista de Fernando Cruz, se centra en lo “no periódica”, pues resalta la característica de “presencia inadvertida en el tiempo” (Cruz Quintana, 2015, p. 305) que tiene el libro.

Otra definición, se encuentra en el diccionario de la Real Academia Española, que se refiere al libro como:

Obra científica, literaria o de cualquier otra índole con extensión suficiente para formar volumen, que puede aparecer impresa o en otro soporte. (Definición de ASALE & RAE, 2020b).

Con esta descripción, el concepto de libro trata de abarcar más allá del libro impreso pues el uso de la frase “o en otro soporte”, aunque poco precisa, abre la posibilidad de integrar al libro digital. Sin embargo, también vale la pena resaltar el uso de la palabra “obra”, que de entrada se refiere a algo producido por el hombre, que en este caso se trata de una obra intelectual. De ahí que, con este concepto, se abre la posibilidad para remarcar que el libro es un producto, concepto que se verá más adelante.

Además, para entender el concepto de libro como objeto, también es necesario estudiar al libro por su historia y en ese sentido, Jeffrey T. Schnapp, reconocido historiador del libro, refiere que los libros “han sido muchas cosas a lo largo de su historia” (Schnapp & Raven, 2020, p. 369) sin embargo, atribuye que los cambios o formas que este ha tenido están relacionados con los cambios en las comunicaciones, que han dado lugar “a nuevos tipos de libros”(Schnapp & Raven, 2020, p. 369).

Cabe señalar que esta investigación no pretende ahondar en la historia del libro, sin embargo, es importante mencionar que de acuerdo con Frédéric Barbier, historiador del libro, el origen del libro es más antiguo que su nombre y surge en la época más primitiva de la humanidad, para expresar ideas, primero con gestos y luego con el lenguaje, como libro oral (Barbier, 2015), el cual se fue transformado con el tiempo y por el desarrollo de tecnologías, como el papel, pues lo que fue pictórico y plasmado en piedra, posteriormente se llevó al papiro, pasando de tecnología en tecnología para la comunicación de ideas, pensamientos y conocimientos, del libro escrito, al impreso y a lo que ahora conocemos como libro digital.

Más adelante se hablará con detalle de los antecedentes del libro digital, por lo pronto, lo interesante es resaltar la visión con la que Schnapp y Barbier, presenta al objeto libro, pues

ambos coinciden en que se trata de producto de la tecnología y que durante toda su historia se ha relacionado con los medios de comunicación, desde el origen de la humanidad y antes de la existencia de la palabra libro y desde luego, desde antes de la existencia del libro digital. De ahí que, nuevamente el concepto de producto se ve dibujado ligeramente, pues de acuerdo con estos autores el libro ha variado en su producción o forma respecto a la tecnología empleada en cada momento de su historia.

Esto también concuerda con la perspectiva de Fernando Cruz, investigador del libro en México, quién menciona que para referirse al concepto de libro es posible hacerlo como un “objeto/tecnología” (Cruz Quintana, 2015) y además señala, que para estudiar al libro es necesario contemplar dos factores esenciales: su función y sus cualidades materiales.

Respecto a la función del libro, ya se refirió como un medio de comunicación o transmisor de la cultura, pues sus páginas puede contener mitos, normas religiosas, cosmogonías y constructos que podrían referirse a la ciencia, la tecnología, el arte, la literatura y todo lo que el hombre ha documentado durante su estancia en la tierra (Díez-Borque, 1985). De ahí que, también es un elemento documental y como tal, de acuerdo con la UNESCO, se compone de dos propiedades importantes, un contenido informativo y su consignación en un soporte material (UNESCO, 2002, p. 7).

Además, el libro es parte del patrimonio cultural de las naciones, ya que su valor no está sólo en el hecho de ser antiguo, pues como ya se dijo en todo momento se constituye como un elemento documental para representar una visión del mundo, que se manifiesta tanto en su contenido informativo como en el soporte material con el que se elabora, de ahí que:

[...] refleja de modo singular un periodo de cambios decisivos en la situación mundial o constituye una contribución destacada a nuestra comprensión del mundo en un momento particularmente trascendente (UNESCO, 2002, p. 17).

Además, su contenido y producción, desde su origen en la historia del hombre, ha generado un impacto social, de ahí que su función históricamente no siempre ha sido la misma, ya que en algunos momentos pudo ser inofensivo, otras revolucionario o incluso considerarse como el “armas más peligrosas que el hombre pudiera desarrollar” (Cruz Quintana, 2015). De ahí que, Escarpit, reconocido escritor francés, refuerza la idea de que el libro debe analizarse tanto en su contenido intelectual como en su forma material y además resalta, que su función comunicadora está dada por ambos elementos, de ahí que, en sus palabras:

[...] un libro no es un objeto como los demás. En la mano, no es sino papel; y el papel no es el libro. Y, sin embargo, también está el libro en las páginas; el pensamiento sólo, sin las palabras impresas, no formaría un libro [...] Un libro se vende, se compra, se cambia [...] Es el fruto de determinadas técnicas que se han puesto al servicio de determinadas intenciones y que permite determinadas utilidades (Escarpit citado por Díez-Borque, 1985, p. 13).

Con lo anterior, Escarpit, retoma el concepto de “obra” que se mencionó al inicio, pues habla de la producción del libro y de que ésta se dirige por una intención y uso específico,

recalcando que, desde su punto de vista, la forma en que se produce y el uso o fin para el que se crea, también forman parte de la definición y función del objeto libro, además del contenido.

En este mismo sentido, también Thomas R. Adams y Nicolas Barker refieren que la forma o apariencia externa e interna (iconotextualidad), durabilidad y tamaño del libro son parte de su función, pues transmiten sensaciones al lector, en donde el libro ejerce un poder sobre quienes lo manejan (Citado por Finkelstein et al., 2006, p. 49) y además definen al libro como un documento bibliográfico, pues, se trata de:

[...]escrito en múltiples copias que su agente, ya sea autor, librero, impresor o editor, o cualquier combinación de estos, produce para consumo público [...] (Citado por Finkelstein et al., 2006, p. 51)

Esto concuerda con lo señalado por Gil & Rodríguez, cuando refieren que el libro desde su creación está determinado para el consumo público, como un producto que se distribuye o vende y se adquiere o compra, de ahí que el libro también va tener una función comercial y ese sentido representa la perspectiva de la industria editorial, en donde el libro es un producto que obedece a la oferta<sup>1</sup> y la demanda<sup>2</sup> (Samuelson et al., 2001, pp. 41-56), por lo que atiende un mercado y satisface ciertas necesidades, de ahí que tiene una cadena de valor que encierra costos de producción, almacenamiento, comercialización y distribución (Gil & Rodríguez, 2011) independientemente del que el producto se genere para comunidades científicas o de investigación que podrían ser muy pequeñas.

Cabe aclarar que la visión del libro como producto desde el punto de vista económico, no sólo obedece a la obtención de una ganancia monetaria pues en el caso de las universidades y gobierno, el beneficio puede ser la satisfacción de las necesidades del mercado o de sus comunidades, de ahí que, también se refiere a la satisfacción de la demanda de la producción de conocimiento (Samuelson et al., 2001, pp. 41-56).

Esta visión del objeto libro como un producto generado por los editores, en el caso del libro digital se ve afectada fuertemente por los avances tecnológicos (Samuelson et al., 2001, pp. 41-56), de ahí que los mismos editores hablen de un cambio disruptivo en la cadena de valor, ya que es necesario incluir la fabricación de software y el uso de medios electromagnéticos y virtuales para la distribución (Piedras, 2013, p. 35), sin embargo este

---

<sup>1</sup> La oferta en términos económicos se desprende de los costos de producción que están determinados por los precios de los factores de producción, los avances tecnológicos, los bienes relacionados y la política económica del gobierno. Los factores de producción son las maquinarias y el trabajo que se necesita para generar un producto; los avances tecnológicos son los que permiten reducir la cantidad de factores para obtener la misma producción; los bienes relacionados son los que podrían sustituir el producto del que se trata y las políticas económicas se refieren a los impuestos y legislaciones aplicadas a un producto. De todo esto se deriva el precio de un producto con el cual es ofertado y se le llama oferta a la producción generada para satisfacer la demanda.

<sup>2</sup> La demanda en términos económicos obedece al ingreso de los consumidores, la dimensión de mercado o población objetivo, los precios de los insumos relacionados con el producto, los gustos y preferencias de los consumidores y los elementos especiales que hacen más necesario un producto en ciertas zonas o para ciertas poblaciones. De ahí que la necesidad de un producto en una comunidad se le llame demanda y esta se satisface con la oferta.

asunto se tratará en el siguiente capítulo con detalle, por lo que sólo vale resaltar que la perspectiva del libro como un producto también forma parte de la esencia del objeto libro.

Así pues, con todo lo anterior es posible decir que para fines de esta investigación el objeto libro se identifica como un documento de presencia inadvertida en el tiempo, que soporta un contenido intelectual y visual, transmisor de pensamientos y sensaciones, valioso por su forma material que incluye el contenido iconotextual, además del medio en el que se contiene. Y que es un producto generado con la intervención de varias personas mediante un proceso o cadena de valor, de ahí que, se puede vender y además se consume (Cruz Quintana, 2015, p. 310).

Por lo tanto, el libro digital o electrónico como veremos más adelante, desde su esencia como objeto libro, se ve afectado tanto en su función comunicadora como en su constitución como producto, por la esencia inmaterial que lo identifica como software. De ahí la importancia de entender cuál es su origen, que a continuación se detalla, para posteriormente hablar de su esencia como software y la importancia de la Ingeniería de software en su producción.

## 1.2. ORIGEN DEL LIBRO DIGITAL

Para hablar del origen del libro digital lo primero que se presenta es un resumen general respecto a lo encontrado en cuanto al nacimiento del libro digital. Además, se ha elaborado una cronología que refleja dos elementos importantes en la historia del libro digital: los avances tecnológicos que impactaron en su nacimiento y la sucesión de hitos encaminados a la creación del libro en el ámbito digital. Por último, se hace un análisis de los datos encontrados para establecer el origen del libro digital en el alcance de este trabajo de investigación.

### 1.2.1. VISIÓN GENERAL.

Respecto a la concepción del libro en el ámbito digital, Jeffrey T. Schnapp señala que, “el linaje del libro electrónico” se remonta al movimiento futurista italiano, que entre 1910 y 1920 (Schnapp & Raven, 2020, p. 371) planteó nuevas perspectivas en la configuración del libro. Pues desde su punto de vista, la creación de publicaciones en papel a partir de la fotografía, el electrotipo<sup>3</sup> y la fototelegrafía<sup>4</sup>, estableció un nuevo diseño del libro que permitió una comunicación visual diferente con el lector e impulsó nuevas perspectivas de lectura, más allá de las palabras. De ahí que, para él, el uso de las tecnologías digitales, sólo permitió potencializar la propuesta de esa época, que incorporó bloques de información, colores, páginas con índices de pulgar e íconos en el libro (Schnapp & Raven, 2020, p. 378).

Desde otra perspectiva, Priani y Galina refieren que un hito importante para la configuración del libro digital o electrónico, sucedió en 1945, cuando IBM y el padre Roberto

<sup>3</sup>Composición tipográfica con procedimientos electroquímicos.

<sup>4</sup>Transmisión de imágenes vía ondas electromagnéticas. Origen de la Televisión

Busa, en Italia (Priani Saisó & Galina Russell, 2019, p. 28) comenzaron por primera vez, a codificar un texto en tarjetas perforadas para generar un índice de las obras completas de Santo Tomás y con lo cual, se demostró la utilidad de la computación en otras áreas del conocimiento.

Además, en coincidencia, el año 1945 también es señalado por otros autores entre ellos Reilly y Henke, como un punto de partida conceptual del libro digital (Reilly, 2003, p. 168), ya que gracias a la publicación del artículo *As We May Think* escrito por Vannevar Bush en dicha época, se dibuja por primera vez, al libro en un ámbito virtual, en el que Vannevar describe una máquina o dispositivo imaginario de nombre *Memex* que otorga la posibilidad a cualquier persona de, “almacenar sus libros ... archivarlos, indexarlos y actualizarlos cuando fuera necesario” (Henke, 2012, p. 11).

Adicionalmente, tanto Reilly como Priani y Galina coinciden en que la construcción del primer libro electrónico fue en 1971 con el proyecto Gutenberg<sup>5</sup> (Reilly, 2003, pp. 85-86) (Priani Saisó & Galina Russell, 2019, pp. 29-30), pues en ese momento se crearon los primeros documentos almacenados en una computadora que contuvieron el texto completo de un libro. Y, de acuerdo con Priani y Galina, este proyecto inició con una visión de “tecnología replicadora” (Priani Saisó & Galina Russell, 2019, p. 29) creando archivos de computadora de contenido únicamente textual, generados con procesos de captura o transcripción a partir de los originales en papel.

Estos primeros libros sólo almacenaron la información en formato ASCII<sup>6</sup>(Reilly, 2003, p. 18), un estándar de codificación de archivo de computadora que en ese entonces permitió replicar en la pantalla, sólo 256 caracteres del alfabeto latino, sin la posibilidad de utilizar acentos o caracteres especiales, cambios en color, forma o alineación en el texto, pero con la ventaja de lectura en varias computadoras, posibilitando la transferencia a varias personas gracias al uso de un medio magnético o electrónico. Esto también, gracias a la existencia de los editores de texto<sup>7</sup>(Reilly, 2003, p. 256), sistemas de software cuyas funcionalidades permitieron mostrar y capturar texto ASCII.

Además, Edwin Reilly refiere que el término libro electrónico fue usado por primera vez por Adries Van Dam<sup>8</sup>, de ahí que se considera que él, acuñó la palabra (Reilly, 2003, pp. 85-86)

<sup>5</sup>Gutenberg Project, persiste hasta la actualidad y puede consultarse en:

[http://www.gutenberg.org/wiki/Gutenberg:The\\_History\\_and\\_Philosophy\\_of\\_Project\\_Gutenberg\\_by\\_Michael\\_Hart](http://www.gutenberg.org/wiki/Gutenberg:The_History_and_Philosophy_of_Project_Gutenberg_by_Michael_Hart).

<sup>6</sup> American Standard Code for Information Interchange (ASCII), Código estándar americano para el intercambio de información, desarrollado por Robert Bemer en 1967, capaz de codificar solo 256 caracteres, no incluye símbolos como diéresis y otros signos diacríticos necesarios para la composición del lenguaje natural como el francés, alemán y español entre otros. Este código prácticamente sólo permite la codificación del idioma inglés.

<sup>7</sup> Programa de computadora que permite preparar, editar e imprimir documentos de texto codificados regularmente en ASCII. Arma los contenidos con caracteres monoespaciados de tamaño uniforme. Los editores, en algunos casos, evolucionaron a sistemas de procesamiento de texto. TVEDIT fue un editor de tiempo compartido de la Universidad de Stanford, escrito por Brian Tolliver en 1963, que permitió intercalar texto directamente sin la intervención de comandos, permitiendo el movimiento del curso. Otro editor fue QED escrito en 1965 por Butler, cuyo descendiente es Vi sistema utilizado hasta estos días, en ambientes UNIX. MS-DOS tiene un editor Edit y Microsoft Windows tiene dos editores de texto, Bloc de notas y Wordpad, este último con funciones de procesador de texto.

<sup>8</sup> Experto estadounidense en gráficos por computadora, creador de FRESS sistema para la construcción de hipertexto.

pues reafirmó el uso del concepto en un artículo titulado *Reading and Writing the electronic book* en 1985, con el cual describe el proyecto *Intermedia* (Yankelovich et al., 1985), conocido como un sistema de libro electrónico que puso en práctica elementos de hipertexto y ambientes gráficos para la navegación entre páginas de documentos electrónicos, de ahí que Van Dam es reconocido en la concepción del libro en el ámbito digital (Ardito, 2000).

Por otro lado, Jeffrey T. Schnapp, señala que desde el punto de vista editorial, la influencia de la tecnología en la construcción de libros data de los años 60, sin embargo, fue hasta los años 70 que se evidencia en Norteamérica, debido a la utilización de la computadora como herramienta para la fotocomposición y construcción de libros, pues ésta permitió la impresión de caracteres especiales en los textos (Schnapp & Raven, 2020, pp. 387-388).

De acuerdo con Henke, el primer lector de libros electrónicos se creó en 1968, por Alan Kay, conocido como Dynabook, el cual fue prototipado por el laboratorio LINC de la compañía Palo Alto Research Center, de Xerox PARC, para que “los estudiantes pudieran acceder a los libros de texto [...] usando una computadora que pudieran llevar consigo” (Henke, 2012, p. 11). Al respecto, Henke señala que el Dynabook buscó igualar las tipografías y relaciones de contraste visual y gráficos del libro impreso, para “instanciar”(Citado por Schnapp & Raven, 2020, p. 383) al libro, en lugar de comprarlo y así, hacer copias fácilmente.

Así pues, conviene señalar que de acuerdo con Priani y Galina, el origen del libro digital partió de la necesidad de distribuir y compartir información entre pares, colaboradores o comunidades, con lo cual se generó el desarrollo de nuevas formas y medios de comunicación en el ámbito de las tecnologías de la información, permitiendo que el texto electrónico se constituyera de formas distintas dando paso a nuevas características y posibilidades en el libro (Priani Saisó & Galina Russell, 2019). De ahí que, a continuación, se desglosan cronológicamente primero un conjunto de hitos o avances en la tecnología que impactaron en el desarrollo del libro digital y posteriormente, se mencionan los sucesos identificados en esta investigación, como parte del desarrollo del libro digital.

---

## 1.2.2. CRONOLOGÍA DEL AVANCE TECNOLÓGICO QUE IMPACTÓ EN EL DESARROLLO DEL LIBRO DIGITAL

---

### 1960

En 1960 Douglas Engelbart<sup>9</sup> y Ted Nelson<sup>10</sup>, influenciados por la idea del Memex de Vannevar Bush, desarrollaron los conceptos de hipertexto e hipermedia, términos que se refieren a la creación de vínculos o nodos entre secciones desconectadas de información textual, gráfica o auditiva, contenida en uno o más archivos de computadora, con la

---

<sup>9</sup> Engelbart hizo una demostración del primer sistema de hipertexto colaborativo, distribuido y en pantalla compartida en 1968

<sup>10</sup> Nelson creó un sistema al que llamó Xanadu, con el cual mostro su propuesta.

posibilidad de navegación a partir de una estructura de nodos, dentro del texto y en un conjunto de elementos interconectados (Reilly, 2003, p. 124).

---

#### 1967

En 1967 Nelson y Andries Van Dam pusieron en práctica los conceptos de hipertexto e hipermedia con la creación de HES<sup>11</sup> el primer sistema para la edición y recuperación de archivos de computadora codificados con hipertexto.

Además, en 1967 con William Tunncliffe, surgió el lenguaje de descripción de página o Page Description Language, conocido como PDL, el cual fue la base para la especificación de la estructura de un documento que permitió, años después, la impresión y visualización en los ambientes gráficos de computadora y del cual, parten las ideas empleadas en el lenguaje de marcado como SGML y PostScript (Reilly, 2003, p. 206).

---

#### 1968

En 1968 Engelbart supero al sistema HES, con la propuesta del NLS u On-Line System, que además de incorporar las características de HES, agregó posibilidades colaborativas para la creación de estructuras de hipertexto (Schnapp & Raven, 2020, p. 384).

---

#### 1969

En 1969 en IBM, Charles F. Goldfarb, Ed Mosher y Ray Lorie crearon un Lenguaje de marcado generalizado estándar, conocido como GML que en 1978 se convirtió en el estándar SGML (Standard Generalized Markup Language) ahora ISO 8879: 1986 y predecesor de HTML (Reilly, 2003, p. 123).

Además, en el año '69 Gary Starkweather en un centro de investigación de Xerox en Webster, Nueva York, inventó lo que ahora se conoce como impresora láser. Sin embargo, fue hasta 1971 que XEROX crea la primera versión que podía imprimir dos páginas por segundo a una resolución de 300 dpi y en 1977 comercializó el primer modelo 9700.

En contraparte IBM en 1976, presentó el modelo 3800, capaz de imprimir 20.000 líneas por minuto y Hewlett-Packard presentó en 1984, su impresora LaserJet con una resolución de 300 ppp y un año después Apple lanzó su Apple LaserWriter (Reilly, 2003, p. 152).

---

#### 1972

En 1972 en la compañía Palo Alto Research Center, de Xerox PARC, Ed McCreight, Chuck Thacker, Butler Lampson, Bob Sproull y David Boggs, desarrollando la computadora conocida como ALTO<sup>12</sup>, considerada también el origen de la PC<sup>13</sup> y modelo base de la Apple

---

<sup>11</sup> Posterior a este, Van Dam con otros colegas crearon un sistema FRESS en 1969, una versión mejorada.

<sup>12</sup> Valorada en ese entonces en 32 mil dólares de la cual se donaron 50 a la universidad Stanford, Carnegie-Mellon y MIT

<sup>13</sup> Personal Computer, término usado por primera vez por Hewlett-Packard.

Macintosh de 1984 (Reilly, 2003, p. 200). Esta máquina presentó por primera vez una Interfaz Gráfica de Usuario o GUI<sup>14</sup> con lo cual, inició la historia de la computadora personal.

---

#### 1975

En 1975, se crea el primer procesador de textos para PC, conocido como Electric Pencil, sin embargo, también BRAVO es considerado uno de los primeros procesadores de texto, debido a que, en su momento, fue un editor de texto innovador, principalmente por la interfaz gráfica que utilizó en el sistema Smalltalk de la computadora ALTO (Reilly, 2003, p. 274).

El procesamiento de texto asistido por computadora se refiere a un software que permite la composición de documentos, desde la captura del texto hasta la selección de estilo tipográfico, la colocación de páginas y la actualización de la tabla de contenido, el índice y la bibliografía. En la actualidad hay quienes aseguran que es difícil distinguir entre un procesador de texto y una herramienta de autoedición, sin embargo, el origen de ambas está en los editores de texto (Reilly, 2003, p. 274). En los años 80, WordStar y FullWrite fueron los procesadores de texto más utilizados y actualmente, Microsoft Word es el líder en el mercado.

---

#### 1976

En 1976 John Gaffney con la ayuda de Martin Newell desarrolló lo que fue el origen de PostScript conocido como "JaM" ("John y Martin"). Pero fue hasta 1982 que John Warnock y Charles Geshchke en Adobe Systems, crean PostScript un PDL especificado como un lenguaje de programación optimizado para representar imágenes en monitores, impresoras, grabadoras de películas, fax y fotocopiadoras (Reilly, 2003, p. 206).

PostScript es un PDL que se caracteriza por conservar en pantalla el formato diseñado para impresión en papel, de ahí que, lo que se ve es lo que se obtiene, esta idea se conoce como WYSIWYG<sup>15</sup>(Reilly, 2003, p. 275) pues permite conservar paginación, encabezados, fuentes tipográficas, además de las metáforas de tabla de contenido, anotaciones, resaltado, marcadores e índices para un documento o libro impreso.

---

#### 1979

La primera computadora portátil o laptop se conoce como Grid Compass<sup>16</sup> (Reilly, 2003, p. 150) diseñada en 1979, sin embargo, las primeras computadoras portátiles compatibles con

---

<sup>14</sup> Presenta en la pantalla del monitor pequeños iconos, barras de herramientas, ventanas, texto e imágenes que dependen del uso de un dispositivo señalador o ratón, que se mueve con la mano para colocar un cursor en un punto elegido de la pantalla y así ejecutar acciones en la computadora.

<sup>15</sup> WYSIWYG (What You See Is What You Get). Software de computadora que con una interfaz gráfica y un apuntador dirigido por un ratón como dispositivo señalador, muestra un facsímil en pantalla construido como un mapa de bits que presenta la apariencia final que tendrá el documento al ser impreso, construido con texto y gráficos.

<sup>16</sup> Fue diseñada por William Moggridge, para Grid Systems Corporation. La Grid Compass pesaba una quinta parte del peso de una PC y tenía un rendimiento equivalente. Se utilizó por la NASA en su programa de transbordadores espaciales de principios de los 80's, tenía

la PC de IBM aparecieron a principios de los años 90, como equipos livianos con todas las capacidades de procesamiento de una computadora personal, de ahí que estos equipos comenzaron a facilitar la movilidad de sus usuarios ampliando las posibilidades para la lectura de documentos hasta mediados de los 90's.

---

#### DÉCADA DE LOS 80'S

En 1980 Tim Berners-Lee creó Inquire<sup>17</sup> un sistema con lo cual permitió la navegación entre sistemas de archivos de computadoras remotas, por lo que es considerado el inicio de lo que ahora se conoce como WWW<sup>18</sup>, sin embargo este concepto no se cristalizó hasta años después con el nacimiento de la internet (Reilly, 2003, pp. 33, 142,124,168 y 275).

En 1984 se desarrolla el primer software de autoedición conocido como PageMaker, por Paul Brainerd de la empresa Aldus. La autoedición o DTP (Desktop Publishing) se refiere a la creación y eventual impresión de documentos de calidad editorial utilizando software por usuarios domésticos y/o de oficina, que podrían o no tener una formación profesional en composición editorial. Otras herramientas de este tipo son QuarkXPress, CorelDraw, Microsoft Publisher y Adobe Indesign (Reilly, 2003, p. 75).

En la década de los 80's se utilizó por primera vez el término cliente-servidor para hacer referencia a redes de tiempo compartido que, a diferencia de los antiguos servidores mainframe, los clientes dejan de ser "terminales tontas" por lo que el cliente y el servidor intercambian roles de manera automática y distribuyen la carga de trabajo durante la ejecución de los procesos, haciendo más eficiente la respuesta del servidor. Este concepto se utilizó más en los inicios de la internet y en la actualidad, sigue siendo de utilidad para el desarrollo de aplicaciones a pesar de que han surgido nuevas formas de arquitectura para el trabajo en red (Reilly, 2003, p. 45).

---

#### DÉCADA DE LOS 90'S

En 1990 Tim Berners-Lee crea HTML o HyperText Markup Language, un lenguaje de descripción de página ideado para la composición y visualización de páginas en la WWW. Este lenguaje es un subconjunto del lenguaje SGML y se expresa en formato de caracteres ASCII estándar, de ahí que, puede crearse y editarse con un simple editor de texto, pero requiere ser interpretado y ejecutado por un navegador Web (Reilly, 2003, pp. 123 y 278) .

---

una memoria de 340 KB (kilobyte) y una pantalla de gráficos electroluminiscentes. Por otro lado, Manny Fernández, quien fundó Gavilan Computer Corporation, en 1984 promovió un equipo como la primera computadora "portátil". Esta se basó en el procesador Intel 8008 con una pantalla de cristal líquido (LCD) de ocho líneas, un panel táctil y una impresora que se adjuntaba a la parte posterior. Las primeras computadoras portátiles IBM compatibles con PC aparecieron hasta principios de la década de 1990.

<sup>17</sup> El sistema *Inquire for Norsk Data Machines*, de acuerdo con Tim fue la abreviatura de *Inquire Inside upon Everything*, un libro que leyó de niño.

<sup>18</sup> Acrónimo de World Wide Web. Cabe señalar que después del trabajo de Berners-Lee, tomó diez años para que CERN viera la utilidad de la implementación de esta idea, que hace uso del hipertexto e hipermedia creados por Nelson y trabajados por él mismo, Engelbart y Van Dam 20 años antes.

En 1991 nace la Internet<sup>19</sup>, pero es hasta 1994 que con la creación del primer Browser<sup>20</sup> o sistema de navegación Web (Reilly, 2003, pp. 33, 142,124,168 y 275) se ven concretados y en su máxima culminación, tanto el concepto WWW como la propia internet. Además, en esa misma época y gracias al desarrollo de dichas tecnologías, nació el concepto de ciberespacio<sup>21</sup>, para referirse a un ambiente intangible o no corpóreo, en donde ocurren fenómenos del pensamiento y la inteligencia (Reilly, 2003, p. 142), en el cual se comenzó a generar una nueva forma de interactuar entre las personas e impulsó una nueva manera de construir documentos para compartirlos desde la comodidad de una computadora.

En 1994 surge el World Wide Web Consortium o W3C<sup>22</sup>, para establecer los protocolos para el intercambio de información en la Web dirigido por Tim Berners-Lee.

En 1997 Tim Bray crea XML o eXtensible Markup Language, el cual se considera un PDL mejorado de HTML, por ser extensible, ya que permite definir etiquetas propias especializadas y enlaces a archivos o entidades no previstas en la propia definición del lenguaje. Cabe señalar que la especificación de XML fue editada y desarrollada por la W3C, e introducido por Jon Bosak (Reilly, 2003, pp. 123 y 278).

Para 1998 surgió el concepto mobile computer<sup>23</sup> (Reilly, 2003, p. 172) para referirse a una computadora con la capacidad de conectarse a Internet de forma inalámbrica que como se verá más adelante, fue importante para el desarrollo de los lectores de libros electrónicos que ahora existen. Además, en el mismo año, el formato Unicode<sup>24</sup> contaba ya con la codificación de 38,887 caracteres de los alfabetos más comunes del mundo, por lo que su uso en internet comenzó a sustituir al formato ASCII, popularizándose para después

<sup>19</sup>La conexión entre computadoras de diversas compañías o entidades comenzó gracias a los acuerdos para la creación de ARPANET entre la agencia de defensa de los EUA y particulares, que fue posible gracias a la invención de: la conmutación de paquetes de Paul Baran de Bolt, Beranek y Newman en los Estados Unidos y Donald Davies en Inglaterra; por la definición del protocolo de comunicaciones TCP / IP de Vinton Cerf y Robert Kahn; la implementación de la interfaz de procesamiento de mensajes IMP de Frank Heart y Leonard Kleinrock; el sistema de nombres de dominio o DNS que junto con los localizadores uniformes de recursos o URL, las direcciones de correo electrónico y el Protocolo de Internet o IP que fueron ideados por Keith Uncapher y Paul Mockapetris.

<sup>20</sup>Los primeros navegadores eran editores de texto, pero los actuales adoptan la forma de una interfaz gráfica de usuario, en algunos casos, subordinada a un sistema operativo por ejemplo Microsoft Edge de Microsoft Windows. Se considera el primero por su popularidad, sin embargo, hubo al menos 6 creados con anterioridad: Inquire, OTHerwise, ViolaWWW, Samba y Midas.

<sup>21</sup>William Gibson en 1984 en su novela *Neuromancer* describió el *ciberespacio* como un entorno artificial creado y mantenido por computadoras. También se usa como sinónimo de supercarretera de la información término popularizado en 1978 por el entonces congresista en los Estados Unidos, Albert Gore, quien se refirió a *un sistema unificado e interactivo de comunicación electrónica*.

<sup>22</sup>Berners-Lee fundó el Consorcio World Wide Web (W3C) en el Laboratorio de Ciencias Informática del Instituto de Tecnología de Massachusetts [MIT/LCS], en colaboración con Organización Europea para la Investigación Nuclear (CERN) y posteriormente se incorporaron, la Universidad de Keio de Japón y el Consorcio Europeo para la Investigación en Informática y Matemáticas (ERCIM). Actualmente se conforma por 441 miembros de todo el mundo. La misión de la W3C es llevar a su máximo potencial la WWW mediante el desarrollo de protocolos y pautas que garanticen su crecimiento a largo plazo, a partir de generar estándares abiertos, criterios de diseño que propicien una "Web para todos" y una "Web en todo", con una visión que contempla una "Web para una interacción enriquecida", una "Web de datos y servicios" y una "Web de confianza".

<sup>23</sup>El concepto se diseñó para lograr que una computadora de mano equipada con un dispositivo inalámbrico pudiera establecer una conexión a Internet. Esta tecnología combina las características de una computadora de mano o portátil, un teléfono celular y un dispositivo de geolocalización o GPS. El primer equipo lo hizo Qualcomm en 1998, sin embargo, resultó demasiado pesada e impráctica esa primera vez.

<sup>24</sup>Desarrollado en 1988 por un consorcio de empresas informáticas, es un código para representar caracteres y su tamaño es de 16 bits, por lo que tiene la capacidad de representar hasta 65.536 caracteres en comparación con el código ASCII que sólo puede representar 256.

convertirse en un estándar para la codificación de documentos en internet (Reilly, 2003, p. 265).

En los 90's Apple presentó su primera computadora PDA que introdujo verdaderamente el concepto en el mercado de handheld computer<sup>25</sup> (Reilly, 2003, p. 117) o computadora de mano y/o computadora con lápiz (Reilly, 2003, p. 201), sin embargo, fue el modelo Palm Pilot<sup>26</sup> que el año 2001 tuvo mayor éxito.

---

### 1.2.3. CRONOLOGÍA DEL DESARROLLO DEL LIBRO DIGITAL

Los avances tecnológicos hasta ahora señalados permitieron o influyeron de una forma directa o indirecta en la construcción de libros en el ámbito digital, sin embargo, además existen momentos que han marcado cambios o avances importantes directamente en su desarrollo, que ahí que, a continuación, se presenta una cronología del desarrollo del libro digital que complementa los acontecimientos señalados en la visión general del inicio de este apartado.

---

#### 1949

En 1949 de acuerdo con Reilly, Edmund Berkeley<sup>27</sup> describió un dispositivo imaginario para la lectura de libros, al que llamó SIMON<sup>28</sup>, que años después inspiró a Alan Kay para desarrollar el primer lector de libros electrónicos, conocido como Dynabook<sup>29</sup> (Reilly, 2003, p. 200).

---

#### 1968

En 1968 Donald Knuth<sup>30</sup>, comienza a escribir el sistema TEX<sup>31</sup> antecesor de LATEX<sup>32</sup>, un lenguaje para la composición tipográfica de fórmulas matemáticas con la posibilidad de

---

<sup>25</sup> Se puede decir que la primera computadora de mano fue la calculadora programable y recibe este nombre porque literalmente es una computadora que se puede sostenerse en la mano. Sin embargo, el primer equipo con lápiz fue la GRIDPAD, diseñada en 1989 por Jeff Hawkins que medía 9 x 12 x 1,4 pulgadas, con un procesador Intel 80 x 86 de 10 MHz y utilizaba MS-DOS además de un monitor de 640 X 400 píxeles, con tarjetas de memoria RAM de 256 o 512 KB. Este equipo se vendía en 2,370 dólares, sin software.

<sup>26</sup> En 1994 Jeff Hawkins y Donna Dubinsky de la empresa Palm Computing presentaron Palm Pilot, una computadora que cabía en el bolsillo de una camisa que tuvo un gran éxito. Estos equipos incluyeron calendarios de citas mensuales y diarias, formularios de notas y una pantalla de calculadora, además de la posibilidad de sincronizar datos con una computadora personal (PC). Los equipos compatibles con PC utilizaron un sistema operativo Microsoft PocketPC 2002 o PalmOS.

<sup>27</sup> Informático estadounidense y escritor del clásico *Giant Brains or Machines That Think*, en el cual describió los circuitos lógicos de una máquina que se podría tener en casa.

<sup>28</sup> Es el nombre que Edmund Berkeley dio a la computadora digital pequeña y programable descrita en su libro clásico *Giant Brains or Machines That Think*.

<sup>29</sup> Considerado el primer prototipo de lector de libros electrónicos desarrollado en 1968, cabe señalar que, además, existió en 1989 el Dynabook de Toshiba, un modelo de computadora personal, que no se relaciona del todo con el Dynabook de Alan Kay.

<sup>30</sup> Donald Knuth fue profesor de la Universidad de Stanford, reconocido especialista en informática y computación.

<sup>31</sup> TEX es un sistema de composición tipográfica creado por Donald Knuth y su nombre deriva de la raíz griega de palabras en inglés "técnica" y "tecnología". TEX es de dominio público y sus algoritmos pueden incorporarse libremente a otros sistemas, pues es un software de código abierto, sin embargo, su creador Knuth condicionó que la palabra TEX se use en el nombre de sus derivados siempre que sean totalmente compatibles con TEX.

<sup>32</sup> LATEX, creado por Leslie Lamport y es un derivado de TEX, permite crear documentos con estructura anidada: capítulos, secciones, subsecciones, ilustraciones y ecuaciones. Además, permite asignar números a listas para generar automáticamente índices y tablas de contenido. Muchos matemáticos y científicos utilizan LATEX para preparar sus artículos antes de enviar a las revistas.

dirigir su salida a una impresora. Este lenguaje lo usaron editores como Addison-Wesley para la impresión de los libros de texto de Knuth, materiales con lenguaje matemático, que fueron una iniciativa para la construcción de libros y por lo cual Donald Knuth fue galardonado con la medalla nacional de Ciencia en EUA en 1979 y reconocido por la ACM<sup>33</sup>(Reilly, 2003, p. 255).

---

#### 1979

En 1979, surge la visión de una biblioteca doméstica con Nicholas Negroponte<sup>34</sup>, quien, conceptualizó un entorno llamado Dataland<sup>35</sup> como una biblioteca conformada con una serie de dispositivos instrumentados alrededor de un cómodo sillón que permitían la navegación en interfaces sonoras, visuales y táctiles (Schnapp & Raven, 2020, p. 386).

Cabe señalar que de acuerdo con un artículo, publicado por Thomas Houston en el 2012 "DataLand" también influyó directamente en los creadores de LISA (Houston, 2012), la primera computadora personal de Apple, con la cual además se superó el alcance del modelo ALTO pues su interfaz gráfica incluyó un sistema operativo que soportaba la ejecución de múltiples tareas.

---

#### 1985

En 1985 los Laboratorios Bellcore desarrollaron un software llamado, Superbook, el cual se considera el primer software que propone un incremento en las posibilidades de lectura de un libro digital, pues incorporó búsquedas en el contenido, gracias al uso del estándar SGML<sup>36</sup>. El proyecto Superbook se diseñó con una arquitectura cliente/servidor<sup>37</sup> y su interfaz de usuario mostraba el título del libro, la tabla de contenido, las páginas y la posibilidad de hacer búsquedas, sin embargo nunca se comercializó (Henke, 2012, p. 12).

---

#### 1987

En 1987 IBM comercializó el primer software para lectura de libros, conocido como BookManager, el cual permitió al usuario construir bibliotecas de libros con la posibilidad de formatear y adaptar el contenido del libro al tamaño de la pantalla de forma automática, además de admitir metáforas derivadas del libro en papel, como tablas de contenido o índice, encabezados, pies de página, números de página, marcadores, anotaciones, estanterías y bibliotecas de libros (Henke, 2012, p. 15).

Además, en los años 80 desde el punto de vista de Jeffrey T. Schnapp, en el ámbito editorial se generó un “diálogo entre el papel y los libros eléctricos”, además de “nuevas expresiones

---

<sup>33</sup> Association for Computing Machinery (ACM).

<sup>34</sup> Nicholas Negroponte cofundó el MIT Media Lab con Jerome B. Wiesner en 1985 y fue su primer director. Negroponte fue un pionero en el campo del diseño asistido por computadora.

<sup>35</sup> Spatial Data Management System o SDMS también conocido como Dataland u oficina del futuro, desarrollado por Architecture Machine Group en el MIT.

<sup>36</sup> Ver año 1969, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>37</sup> Ver década de los 80'S, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

de afición y cultura libresca” (Schnapp & Raven, 2020, pp. 387-388), pues la industrialización de la computadora personal, periféricos (impresión láser<sup>38</sup>) y el desarrollo de software (procesador de textos<sup>39</sup>, herramientas de autoedición<sup>40</sup> y los lenguajes PDL<sup>41</sup> y PostScript<sup>42</sup>) permitió la composición de libros desde casa, con la selección de estilos tipográficos de diversos colores, formas, tamaños y orientación; control de palabras, líneas y espaciado que facilitó la paginación y construcción de tablas de contenido, índices y bibliografías, haciendo posible la composición completa de un libro impreso en una computadora personal.

---

#### 1991

En 1991 Computer Literacy una librería de Silicon Valley en Estados Unidos, comenzó a vender libros impresos por internet (*Amazon.Com / History & Facts*, s. f.), sin embargo, Paul Ceruzzi menciona que es hasta julio de 1995, que Jeff Bezos abre una librería en línea conocida ahora como amazon.com. Amazon fue innovadora desde sus inicios por las dinámicas que estableció con los lectores, pues en su sitio se animaba a los clientes a escribir reseñas de los libros adquiridos y entre ellos era posible calificar las reseñas, generando intercambio entre la comunidad de lectores, esto además generó que estudiantes miraran las reseñas para referirse a los libros con sus profesores a pesar de no haber leído los materiales y lo más interesante, los autores empezaron a visitar el sitio de Amazon para saber qué tan bien o mal, eran recibidas sus obras (Ceruzzi et al., 2003, pp. 326-328).

---

#### 1993

De acuerdo con Feldman, Keep & McLaughlin, de 1990 a 1993 Sony comercializó los primeros lectores de libros: Sony Data Discman y Bookman, sin embargo ellos afirman que “no tuvieron éxito” (Feldman, Keep & McLaughlin citado por Henke, 2012, p. 12). Estos equipos fueron la primera aproximación de lo que ahora conocemos como PDA<sup>43</sup> o asistente digital personal y en su momento, se componían de un lector de CD ROM, memoria, pantalla y teclado. Se comercializaron como reproductores de libros y tuvieron un software que permitía crear libros para ser leídos en estos dispositivos.

El software de Discman era propietario de Sony por lo que los libros creados con éste, sólo se podían leer en el dispositivo Discman, en cambio Bookman, usó un software que leía CD-ROM compatibles con MS-DOS.

---

<sup>38</sup> Ver año 1969, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>39</sup> Ver año 1975, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>40</sup> Ver década de los 80'S, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>41</sup> Ver año 1967, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>42</sup> Ver año 1976, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>43</sup> Personal Digital Assistant conocido como Apple Newton, fue una computadora de mano que permitió almacenar, acceder y organizar información. Estos equipos basados en pantallas no tienen teclado físico, por lo que su acceso es escribiendo o tocando en su pantalla con un lápiz, de ahí también, el término computadora con lápiz.

Sin embargo, de acuerdo con Henke, ambos lectores tuvieron problemas con la pantalla, pues dificultaban la lectura ya que sólo mostraba una parte de la página por su tamaño pequeño, la batería era de poca duración y las funciones de búsqueda deficientes, además los dispositivos no podía comunicarse con otras computadoras por lo que esto, su alto costo y las limitaciones de contenidos generaron su salida del mercado, a pesar de que al momento de su comercialización se crearon 20 títulos distribuidos para este lector, entre los cuales estuvieron el Diccionario enciclopédico de Hutchinson, Biblioteca del Futuro y el Diccionario y tesoro de Oxford. (Henke, 2012, pp. 12-13).

---

#### 1994

En 1994, por otro lado, la compañía Adobe comercializó por primera vez Acrobat Reader, un software que permitió leer libros electrónicos, en formato PDF.

El formato PDF se basa en el lenguaje PostScript e incorpora posibilidades que no existen en los libros en papel, como vistas en miniatura de las páginas, funciones de búsqueda en el texto completo, firmas digitales, visualización y navegación de página en página o a dos páginas, simulando el libro abierto para la lectura en pantalla.

---

#### 1996

En 1996 la digitalización de medios impresos y cintas con audio y video se aceleró, esto también se acompañó de la producción, redacción y diseño de estos recursos directamente en la computadora, convergiendo en lo que ahora se conoce como multimedia, de ahí que en 1997 la Organización Internacional del Trabajo (OIT) llevó a cabo el primer Simposio sobre Convergencia Multimedia (Lebert, 2010).

---

#### 1997

En 1997 investigadores del Media Lab del MIT<sup>44</sup> crean E-Ink, una compañía que desarrolló la tecnología que permite a las pantallas de los dispositivos PDA, asemejar la textura y características del papel, para que, a pesar de los cambios en la intensidad de la luz permitan la lectura.

La tecnología de E-Ink se basa en el LCD<sup>45</sup> (Reilly, 2003, p. 156) adaptado a un tamaño pequeño y mejorado, conocido como TFT LCD<sup>46</sup> y se ha comercializado como tecnología ePaper<sup>47</sup>(Rus, 2021) o de tinta electrónica, sin embargo estas pantallas se presentaron por

---

<sup>44</sup> Massachusetts Institute of Technology

<sup>45</sup> Liquid Crystal Display o pantalla de cristal líquido, fue inventada por James Ferguson en 1979 y es el tipo de pantalla plana más común o flat-panel display (FPD). Utiliza una forma inusual de la materia descubierta en 1888 y conocida como cristal líquido. Un cristal líquido, se comporta mecánicamente como un líquido, pero ópticamente como un sólido y es tan sensibles que responden al calor, la luz, el sonido, la presión mecánica, los campos electromagnéticos, la radiación y a ciertos vapores químicos. Estas pantallas comenzaron a desplazar al monitor CRT (cathode ray tube) o de tubos de rayos catódicos, usadas en los años 80's y 90's.

<sup>46</sup> El TFT LCD o Thin-film transistor LCD es una pantalla LCD de transistores de película delgada.

<sup>47</sup> La compañía E-Ink genera las pantallas y las vende a los proveedores de PDA más importantes del mundo, su tecnología se basa en EPD (Electronic Paper Display). El EPD Es de bajo consumo de energía, pues sólo utiliza energía cuando el contenido se actualiza o cambia. Es perfectamente legible incluso con luz solar directa, pero en lugares poco iluminados puede ser una problemática,

primera vez como un prototipo para los lectores de libros electrónicos hasta 2002 (Lebert, 2010).

---

#### 1998

En 1998, comenzó la comercialización del PDA Rocket eBook, que fue diseñado por NuvoMedia para “imitar al libro de bolsillo tanto en forma como en función” (McCusker citado por Henke, 2012, p. 13) con una pantalla que permitía cambiar de página al contacto de un sólo un botón y opciones de menú que se mostraban en la misma pantalla para marcar páginas y cambiar tamaños de letra.

---

#### 1999

En 1999 aparece, la biblioteca municipal de Helsinki de Finlandia, como la primera biblioteca con servicios en la web. Además, surge el primer catálogo de recursos lingüísticos Human-Languages Page, con más de mil ochocientos recursos, ahora conocido como [ilovelanguages.org](http://ilovelanguages.org). También en el mismo año se crea la biblioteca digital multilingüe Athena, la National Academy Press puso en acceso abierto, cientos de libros en formato de texto y Palm Digital Media logró colocar más de dos mil títulos en su sitio web [Peanutpress.com](http://Peanutpress.com).

En este mismo año, Adobe introdujo Web Buy con Acrobat Reader, otro software, diseñado para la compra de libros electrónicos en formato PDF que además permitió otorgar protección de derechos a autores y editores (Henke, 2012, p. 15).

También en 1999 Microsoft comenzó a comercializar Microsoft Reader (Henke, 2012, p. 16), etiquetado como un software de lectura de libros para PDA con sistema operativo Pocket PC de Microsoft, que en el año 2000 se distribuyó para computadoras con sistema operativo Windows de Microsoft y mejoró la visualización en pantallas LCD, con funciones de búsqueda en colecciones de libros e incorporó DRM<sup>48</sup>(«Sistemas DRM», 2018) para evitar que los libros se lean sin comprarlos.

Además, en el 99 surgió OeB u Open eBook, el primer formato estandarizado para el libro digital que en el año 2005 la IDPF<sup>49</sup> y la W3C, transformaron en el formato conocido como ePub<sup>50</sup>. Un formato estándar basado en XML y diseñado para la construcción de libros digitales que permite facilitar la diagramación del contenido con la posibilidad de que el texto se ajuste a cualquier tamaño de pantalla (Lebert, 2010).

---

inicialmente sólo eran en blanco y negro o con un tercer color adicional (rojo o amarillo), sin embargo, ahora ya manejan hasta 4,096 colores.

<sup>48</sup> Digital Rights Management, permite el empaquetado y encriptación del contenido digital, además de controlar el acceso desde un servidor que tiene un registro del tipo de licenciamiento para determinar el tipo de copia y reproducción pagada por el usuario, requiere de un lector compatible con el formato de archivo, el cual valida el licenciamiento y decodifica el contenido para poder desplegarlo al usuario.

<sup>49</sup> International Digital Publishing Forum o IDPF un consorcio industrial internacional que reúne a constructores, informáticos, editores, libreros y especialistas digitales, el cual surgió como Open eBook Forum.

<sup>50</sup> Estándar OeBPS Open eBook Publication Structure.

Así pues, de acuerdo con Jeffrey T. Schnapp, en la década de los 90's se intensificó la construcción de documentos en el ciberespacio gracias a la creación de HTML<sup>51</sup> y XML<sup>52</sup>, generando un mayor intercambio de contenidos académicos y científicos, además de promover la lectura en diferentes lugares y computadoras. Y en el ámbito editorial, para Schnapp, los cambios se volvieron disruptivos pues:

[...]Los autores se convirtieron cada vez más en procesadores de texto y tipógrafos, a medida que la labor de los profesionales de la composición y el diseño que alguna vez fueron distintivos fue desplazada por componentes integrales de los paquetes de edición de texto [...] los archivos resultantes se compartieron directamente como una operación de publicación y distribución lista para usar. Ya sea impresos físicamente, distribuidos en forma digital para la lectura en pantalla, o producidos en casa o impresos bajo demanda[...] (Schnapp & Raven, 2020, p. 388)

---

#### AÑO 2000

En el año 2000 surgen dos dispositivos lectores de libros electrónicos eBookMan y Gemstar eBook, el primero creado por la empresa Franklin como un dispositivo multimedia, en blanco y negro, que inicialmente utilizó un software lector propietario de Franklin Reader, sin embargo, años después adoptó el formato Mobipocket Reader<sup>53</sup>.

Gemstar eBook sucesor del Rocket eBook<sup>54</sup>, construido y vendido bajo la marca RCA, que perteneció a Thomson Multimedia, tuvo modelo con pantalla en blanco y negro y otro en color, utilizó un formato derivado del OeB<sup>55</sup> a pesar de que el software lector conocido como SoftBook Reader era propietario por lo cual era compatible con ePub, sin embargo, de acuerdo con Lebert, este equipo desapareció en 2003 (Lebert, 2010).

---

#### AÑO 2001

En marzo de 2001, la empresa Mobipocket S.A., crea Mobipocket Reader y con esto, surge un nuevo formato de codificación de libros electrónicos conocido como Mobipocket<sup>56</sup> que inicialmente usó la extensión .prc y posteriormente .mobi, este formato es propietario y codificado en un archivo binario<sup>57</sup> que en 2005 fue adquirido por Amazon<sup>58</sup> (Library of Congress, 2019).

---

<sup>51</sup> Ver década de los 90'S, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>52</sup> Ver década de los 90'S, en cronología del avance tecnológico que impactó en el desarrollo del libro digital.

<sup>53</sup> ver año 2001. Cronología del desarrollo del libro digital.

<sup>54</sup> ver año 1998. Cronología del desarrollo del libro digital.

<sup>55</sup> También conocido como ePub

<sup>56</sup> Este formato de archivo se declara como universal porque su compañía lo introdujo como un formato con la facilidad de leerse en diversos modelos de PDA o lectores de libros y en varios idiomas, además de incluir teléfonos inteligentes de ahí que con su lector Mobipocket Reader se vuelve multiplataformas haciendo que el formato sea compatible con una amplia diversidad de equipos

<sup>57</sup> Se refiere a archivos de computadora cuyo contenido se encuentra codificado en código binario

<sup>58</sup> Amazon utilizó archivos con la extensión .mobi hasta 2017 y después introdujo nuevas herramientas, que no utilizaron el formato .mobi

---

### AÑO 2003

Para el año 2003, como ya se vio, existían en el mercado varios dispositivos de lectura: Gemstar/NuvoMedia Rocket eBook, Softbook, Everybook Journal, eb Dedicated Reader, Millenium eBook Reader y Cybook, además de diverso software multiplataforma para la lectura de libros como Acrobat Reader, Microsoft Reader, Mobipocket Reader y diversos formatos de archivo para la construcción de libros digitales como: mobi, pdf y epub.

De ahí que, de acuerdo con Schnapp & Raven, a principios de este siglo, uno de los principales retos fue mejorar la ergonomía de los PDA o lectores de libros electrónicos para simular o emular las sensaciones y ventajas de lectura en un libro impreso, como el color del papel, la facilidad para el cambio de páginas, la lectura bajo espacios con mucha luz y el ajuste de la pantalla para aproximarse al tamaño de un libro de bolsillo con la capacidad de sostenerlo en una mano (Schnapp & Raven, 2020, p. 328). Pues en contraste, Priani y Galina señalan que para las personas la lectura en pantalla era poco atractiva, más aún cuando se trataba de textos muy largos (Priani Saisó & Galina Russell, 2019, p. 40), de ahí que la búsqueda por mejorar esas sensaciones se convirtió en un reto.

---

### AÑO 2004

En 2004 la empresa Google establece negociaciones con algunos editores, universidades e incluso autores de libros mediante un proyecto que surgió como Google Print, para digitalizar una gran cantidad de contenidos impresos, que en 2005 se suspendió por conflictos de derechos de autor, sin embargo en 2006 Google liberó Google Books con miles de libros en acceso abierto producto del proyecto Google Print, por lo cual en 2008 la empresa enfrentó un conflicto legal que lo obligó a retirar por un tiempo algunos materiales hasta garantizar el respeto a los derechos de los autores, de ahí que años después debió aplicar cambios y restricciones en los materiales disponibles en su plataforma (Lebert, 2010).

---

### AÑOS 2007-2009

Para 2007 Amazon liberó en el mercado su lector conocido como Kindle y la empresa Barnes and Noble en 2009 liberó Nook, ambos equipos móviles con tecnología de tinta electrónica y posicionados fuertemente en el mercado junto con las tabletas iPad de Apple (Schnapp & Raven, 2020, p. 389).

Además, de acuerdo con Schnapp & Raven, en estos años comenzó a cambiar la forma de interactuar con el libro pues se volvió común conformar comunidades virtuales en sitios web<sup>59</sup> y compartir opiniones y sugerencias de lectura, además, con los sistemas de calificación para los libros se “invaden muchas de las funciones que antes realizaban los críticos y las listas de bestsellers” (Schnapp & Raven, 2020, p. 389) y el dispositivo

---

<sup>59</sup> <https://www.goodreads.com/> sitio web

electrónico se ha convertido en “la manera perfecta de llevarse de vacaciones una biblioteca de bestsellers” (Schnapp & Raven, 2020, p. 388).

---

#### 1.2.4. ANÁLISIS DEL ORIGEN DEL LIBRO DIGITAL

De acuerdo con Priani y Galina, el libro electrónico o digital como ahora lo conocemos tiene tres líneas de origen, como texto en el ámbito digital, como formato de codificación y como dispositivo de lectura (Priani Saisó & Galina Russell, 2019, p. 27), lo cual concuerda con todo lo que ya se vio en este recorrido cronológico.

En el mismo sentido Schnapp & Raven, aseguran que el texto en el ámbito digital ha cambiado y esto también ha generado un cambio en las personas, sean autores, editores y lectores pues por un lado los autores se vuelven accesibles, ya que es posible intercambiar ideas con ellos, calificarlos e incluso colaborar con ellos, la lectura se vuelve híbrida pues además es posible hacer anotaciones que se hacen disponibles o incluso colaborativas (Schnapp & Raven, 2020, p. 389).

De ahí que, incluso la creación del libro impreso ha cambiado, ya que en la actualidad un libro nace digitalmente con todas sus características icontextuales e incluso es posible condicionar su impresión a partir de la demanda de lectores, por lo que se vuelve efímera su existencia (Schnapp & Raven, 2020, p. 390).

Por lo tanto, para Schnapp y Raven el origen del libro digital también está dado por los cambios en las formas de edición, publicación, distribución y lectura, de ahí que señalan que un libro o publicación impresa puede clasificarse de acuerdo con el tipo de tecnología empleada en su construcción, por ejemplo: sitio web, producto multimedia, sistema informático, repositorio, base de datos o e-pub, conceptos completamente nuevos en el ámbito del libro, pero además se emplea nuevos conceptos importantes para su construcción como el de *arquitectura de información* (Schnapp & Raven, 2020, p. 390).

Por otro lado, el término libro electrónico, se comenzó a utilizar en los años setenta, sin embargo, Voutssás considera más adecuado ahora usar la palabra digital, pues argumenta que “la estructura fundamental de la información se encuentra en formatos digitales” (Voutssás, 2006, p. 98) por lo que el libro no ocupa un espacio físico y se distribuye por la Web reduciendo el uso de soportes electromagnéticos, con lo cual, el medio de transporte es intangible, altamente volátil además de cambiante y puramente digital.

De ahí que, con todo lo visto hasta el momento, se confirma lo que autores como Priani y Galina mencionan, respecto a que el libro digital ha tenido transformaciones revolucionarias respecto al libro impreso por lo que además, ha afectado también la producción, circulación y comercialización generando nuevas formas de creación, distribución y lectura para el libro en el ámbito digital (Priani Saisó & Galina Russell, 2019).

### 1.3. ESENCIA DEL LIBRO DIGITAL COMO SOFTWARE

Para profundizar en las propiedades del libro digital derivadas de su esencia intangible, es necesario analizar cuáles son las características que éste adquiere, al constituirse como un objeto construido en un ámbito digital.

Un objeto digital de acuerdo con la definición empleada por el Diccionario de Bibliotecología y Ciencias de la Información es:

[...] un tipo de estructura de datos que consiste en contenido digital, un identificador único para el contenido (denominado "identificador") y otros datos sobre el contenido, por ejemplo, metadatos de derechos. (Reitz, 2014)

El uso de la palabra estructura de datos, desde el ámbito computacional se refiere a una colección de valores interconectados que permite operaciones entre ellos. De acuerdo con el tipo de interconexión, las estructuras se pueden clasificar como conjuntos, arreglos, listas, matrices entre otras que pueden ser más complejas.

Cada estructura de datos cuenta con operaciones para lectura y modificación y en algunos casos operaciones sofisticadas como ordenamiento y búsqueda por mencionar algunas (Reilly, 2003, p. 71). Por lo anterior, podemos decir que un objeto digital se constituye como valores agrupados y enlazados, en donde la forma de leer, modificar u operar con dichos valores, depende del tipo de interconexión con la que se construye. Además, se menciona la existencia de datos acerca del contenido o metadatos, de los cuales se hablará más adelante con profundidad, por lo pronto es importante sólo remarcar la importancia de su presencia.

Otra definición es la que hace la organización internacional dedicada a la investigación de la autenticidad de los registros en los sistemas electrónicos, conocida como InterPARES<sup>60</sup>, que describe al objeto digital como:

Una agregación discreta de una o más cadenas de bits y los metadatos acerca de las propiedades del objeto y, si corresponde, los métodos para realizar operaciones en el objeto. (InterPARES 2, 2021)

Este concepto aparentemente coincide con la definición anterior, primero respecto al hecho de que debe contener metadatos y segundo, por la existencia de una estructura de datos, que como ya se vio se compone de métodos que permiten operar con el objeto.

Sin embargo la estructura de datos a la que esta definición se refiere parte de un concepto más general, pues emplea el término bit<sup>61</sup> (Reilly, 2003, p. 31), nombre que recibe un dígito

<sup>60</sup> International Research on Permanent Authentic Records in Electronic Systems

<sup>61</sup> La palabra bit fue acuñada por Claude Shannon como una contracción de "binary digit" en 1936. Esta palabra la uso por primera vez el matemático y estadista John Tukey en 1946.

en el sistema binario<sup>62</sup> y se trata de una representación simbólica del estado de un circuito dentro de la computadora, en donde los bits se encuentran agrupados en arreglos o estructuras de datos de agregación continua, por lo que es este caso, la definición de objeto digital, en esencia expresa que la estructura de datos de un objeto es la que parte de la forma en que se encuentra almacenado en la computadora o hardware.

De lo anterior, que para fines de esta investigación un objeto digital se puede entender como una estructura de datos, con metadatos, constituida mínimamente a partir de arreglos de bits dentro del hardware de la computadora, con un identificador y la posibilidad al menos, de realizar operaciones como lectura y escritura e incluso más complejas, si el tipo de estructura lo permite, de ahí que los elementos identificados como componentes de un objeto digital son:

1. Almacenado en Hardware.
2. Estructura de datos con un Identificador (datos interconectados) y operaciones.
3. Metadatos (Información al respecto).

Hasta aquí, conviene hacer una pausa, para revisar ahora cómo se define al libro electrónico, pues es conveniente determinar si existen o no las características que lo delimiten como objeto digital y para esto la primera descripción seleccionada es la elaborada por la DPC o Digital Preservation Coalition, en donde un libro electrónico es:

[...] un 'análogo digital' de lo que percibimos como un libro [...]Una instancia de un libro de 'papel y tinta' en un medio electrónico (Kirchhoff & Morrissey, 2014, p. 4)

DPC, identifica dos componentes de manera muy diferenciada, un “análogo digital” y el “medio electrónico”, esto también se ve reflejado en las siguientes tres definiciones:

[...] es una publicación libro-enriquecido en formato digital, que consiste en texto, imágenes o ambos, legible en computadoras u otros dispositivos electrónicos [...] (Gardiner & Musto citados por Khalid, 2014, p. 244)

Libro en formato adecuado para leerse en un dispositivo electrónico (ASALE & RAE, 2020a)

[...] Es sólo un archivo almacenado que contiene el texto completo de un libro [...] (Andries van Dam citado por Reilly, 2003, p. 85)

En estas descripciones el “análogo digital” que menciona DPC se detalla un poco más, pues, por un lado, se habla de un “libro-enriquecido”, formato digital en el que el contenido puede ser texto, imágenes o una combinación. En la segunda definición se dice que el formato debe ser el adecuado y la última, habla de que es un archivo. Por lo que, las cuatro descripciones hacen referencia a dispositivos de almacenamiento que sirven como

---

<sup>62</sup> Sistema numérico que matemáticamente se compone sólo de dos dígitos 0 y 1, de tal forma que los números se expresan utilizando secuencias de ceros y unos, agregándose uno detrás de otro de forma consecutiva. En el ámbito de la computación este sistema se emplea para representar simbólicamente los dos estados que un circuito puede asumir, encendido y apagado.

contenedor para la lectura o legibilidad del libro, pero el libro es el elemento “análogo digital” y no el dispositivo.

Sin embargo, también existen definiciones en donde el “análogo digital” del libro es el “medio electrónico”, por ejemplo:

[...] una metáfora o mito generalizado que proyecta una imagen tanto para los diseñadores como para los usuarios de ser como un libro “de papel” [...] (Barker citado por Henke, 2012, p. 19)

[...] dispositivo de lectura que permite al usuario ver colecciones de documentos o texto donde el contenido está representado por un libro [...] (Borcher y Lemken citados por Henke, 2012, p. 18)

son los dispositivos de hardware, como asistentes digitales personales, computadoras portátiles, teléfonos móviles, etc., que se utilizan para acceder a la palabra impresa, ya sea en forma de bases de datos o estructuras (Feldman citado por Henke, 2012, p. 19)

e-book, es la presentación de archivos electrónicos en pantallas digitales [...] dispositivos dedicados a lectura digital (Romano, 2002)

En estos ejemplos se establece al libro, como el hardware de lectura que incluye a las computadoras, portátiles, PDA, teléfonos inteligentes y/o dispositivos diseñados como lectores de documentos y que proyectan una metáfora o representación de la palabra impresa, aunque la metáfora del libro recae sobre el hardware.

Al respecto, Harold Henke, especialista en libros electrónicos, menciona que la forma más simple de referirse al libro electrónico es como un dispositivo de lectura (Feldman citado por Henke, 2012, p. 19), sin embargo es necesario tomar en cuenta que además esta nueva forma del libro también involucra:

“crear un nuevo formato de lectura que reinterprete el libro en papel en un formato nuevo e interactivo” (Feldman citado por Henke, 2012, p. 19)

por eso:

[...]sus páginas [...] están compuestas por [...] páginas reactivas y dinámicas de información multimedia[...] (Barker citado por Henke, 2012, p. 18)

De ahí que, señalarlo sólo como un dispositivo de lectura de forma aislada, desde la perspectiva de Henke, limitan al libro electrónico, pues actualmente la lectura en línea es otra característica que lo distingue del libro impreso (Henke, 2012, p. 19), sin olvidar que el usuario lector siempre “espera contenido bien diseñado (libros) que se pueda usar fácilmente” de ahí que, la “adaptación de metáforas de libros en papel y características similares a las de un libro” (Feldman citado por Henke, 2012, p. 19) persiste o se hace necesaria.

De lo anterior que Henke asegura que los libros electrónicos son “productos [...] tanto de hardware como de software” (Henke, 2012, p. 18). En contraste, Edwin Reilly, doctor en ciencias de la computación, refiere que:

[...] las empresas de software ahora ofrecen programas para computadora personal y portátil [...] que hacen más fácil la navegación a través de los archivos de libros electrónicos, y hay dispositivos de tamaño especial diseñados específicamente para facilitar la lectura de materiales como libros electrónicos (Reilly, 2003, p. 86)

Lo anterior retoma dos perspectivas que en efecto se vieron como parte de la historia del libro electrónico, en la sección anterior de este capítulo, en el que se relata detalladamente cómo se fue desarrollando por un lado el producto que ahora se conoce como lector de libros electrónicos o e-reader, un hardware que ha buscado mejorar principalmente la movilidad, que deriva de la metáfora del libro impreso de bolsillo. Y, por otro lado, los productos de software o programas de navegación y formatos de archivo que de igual manera surgieron para posibilitar primero la captura de contenidos, luego impresión y la composición iconotextual con lo cual, también es posible simular la metáfora de lectura, en donde esta última, agregó nuevas formas y dinámicas gracias al uso del hipertexto, la hipermedia, las búsquedas de texto completo y la organización virtual en estantes y colecciones de libros en el ámbito digital.

Con todo esto es posible identificar elementos en común extraídos de las descripciones y definiciones del libro electrónico, que se vieron anteriormente:

1. Dispositivo electrónico o hardware de lectura o visualización
2. Archivo almacenado (Andries van Dam citado por Reilly, 2003, p. 85)
3. Que contiene la palabra impresa, texto, imágenes o ambos (Gardiner & Musto citados por Khalid, 2014, p. 244)
4. Con formato adecuado, (ASALE & RAE, 2020a)
5. En base de datos o estructura de datos (Feldman citado por Henke, 2012, p. 19).
6. Programas de computadora o software (Reilly, 2003, p. 86)

Antes de analizar estos elementos es conveniente ahondar primero en el término Software, pues como ya se vio, es mencionado por Reilly y Henke de forma coincidente.

La palabra Software se acuñó por primera vez en 1958 por John Wilder Tukey<sup>63</sup> para describir a “la información programática intangible” (Reilly, 2003, p. 237) y hace referencia a los componentes de la computadora que no son hardware (Bourque et al., 2014, p. XVII), Edwin Reilly, doctor en ciencias de la computación, por otra parte menciona que:

---

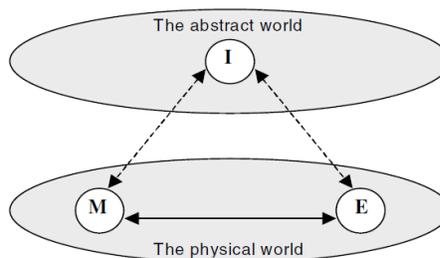
<sup>63</sup> Reconocido estadístico estadounidense e investigador; fundó el departamento de estadísticas de Princeton, fue investigador de los Laboratorios Bell de AT&T y consultor para el Educational Testing Service, Xerox Corp. y Merck & Co.

[...] El software se puede instanciar en hardware, como cuando se graba en un CD-ROM para su distribución, pero el CD no es el software; su información almacenada lo es [...] (Reilly, 2003, p. 238)

Esta analogía de Reilly también nos lleva a reflexionar en el término software dentro del ámbito del libro electrónico, pues el lector del libro electrónico sin el software no podría ser en sí mismo un libro. Sin embargo, leer un libro en una computadora portátil o de escritorio definitivamente no genera la misma sensación o satisfacción, al usuario lector, en comparación a hacerlo con el dispositivo diseñado especialmente para esto.

Lo anterior, se relaciona con el punto de vista de la informática cognitiva<sup>64</sup>, en donde el software como producto tecnológico, se presenta como un modelo abstracto de un objeto del mundo físico (Wang, 2002b) que se construye a partir del modelo Información–Materia–Energía (IME), ver Figura 3, que se explica partiendo de la idea en que los seres humanos viven en dos mundos, un mundo físico o concreto y un mundo abstracto o percibido. En dónde el primero, se construye a partir de materia y energía y el segundo con información que deriva de los dos anteriores (Wang, 2002a).

Figura 3. Visión del mundo en el Modelo IME.



Fuente: (Wang, 2002a).

De ahí que, desde esta perspectiva, entonces, la información abstraída de aspectos del mundo natural que se procesa mentalmente, representa y materializa digitalmente (Wang, 2002a), lo que se conoce como software, definido como:

un tipo especial de información con comportamiento computacional que va a permitir la interacción entre el mundo abstracto y el mundo físico (Wang, 2008, pp. 12-13)

Ahora bien, regresando a los elementos identificados en el libro digital y retomando el hecho de que la palabra software aparece como parte de éstos elementos, se decidió hacer un ejercicio y tomar la frase de la definición de software de Wang, “un tipo especial de información con comportamiento computacional” (Wang, 2008, pp. 12-13) para integrarla en la definición del libro digital de Barker y remplazar la frase “una metáfora o mito generalizado” (Barker citado por Henke, 2012, p. 19) con lo cual se obtuvo la siguiente definición:

<sup>64</sup> La informática cognitiva (IC) es una disciplina que estudia los mecanismos internos de procesamiento de información del cerebro y su aplicación en la ingeniería desde un enfoque interdisciplinario.

*“un tipo especial de información con comportamiento computacional” “que proyecta una imagen tanto para los diseñadores como para los usuarios de ser como un libro “de papel”.*

Este ejercicio nos permite ver que la concepción del libro electrónico, parte de la abstracción de información respecto a la representación del libro impreso como materia y energía, tal como Wang refiere que es el software en nuestro mundo material. De ahí que para entender cómo se compone un libro electrónico, es necesario comprender qué es el software y la relación de este como objeto digital dentro de un hardware, pues este último es el medio por el cual interactúa con el mundo físico en el que vivimos, sin embargo, la composición del libro como objeto digital se deriva de su naturaleza como software.

De lo anterior la importancia de entender desde el ámbito de la ingeniería de software, el término software, pues como veremos más adelante esta área es la responsable del estudio y desarrollo del software, por lo que lo define como:

Programas de computadora, procedimientos y, posiblemente, documentación asociada y datos relacionados con el funcionamiento de un sistema informático<sup>65</sup> (SEVOCAB et al., 2021d)

En esta descripción lo que Wang señaló como “información con comportamiento computacional” y Tukey identificó como “información programática intangible”, aquí se establece como programas de computadoras, que en esencia son secuencias de instrucciones para representar un algoritmo<sup>66</sup> (Viana, 1995) y que se define de acuerdo con el estándar ISO/IEC/IEEE 24765:2017<sup>67</sup>, como:

La combinación de instrucciones de computadora y estructuras de datos que permiten que el hardware de la computadora realice operaciones de control o cómputo (SEVOCAB et al., 2021a).

Sin embargo, de acuerdo con la definición de software, además de los programas de computadora, se debe incluir, la documentación, datos y procedimientos relacionados con el funcionamiento de un sistema informático, este último definido como:

sistema que contiene uno o más componentes y elementos como computadoras (hardware), software asociado y datos (SEVOCAB et al., 2021b)

De lo anterior que el software, se inserta en sistemas informáticos en donde coexiste con elementos de hardware y más software, por lo que también se compone de características no funcionales, como documentación, datos y procedimientos relacionados como la arquitectura del sistema informático en el que cohabita, al respecto, Wang señala que el software tiene un comportamiento que puede ser dinámico o estático, en donde el primero

---

<sup>65</sup> Estándar ISO / IEC 19770-1: 2017 e ISO/IEC/IEEE 26513:2017 Systems and software engineering--Requirements for testers and reviewers of information for users.

<sup>66</sup> Emplea un lenguaje especial para indicar a la máquina las acciones que debía ejecutar. Los primeros algoritmos se codificaron en un inicio a través de cintas y tarjetas perforadas.

<sup>67</sup> Systems and software engineering-Vocabulary (Vocabulario de Ingeniería de sistemas y software)

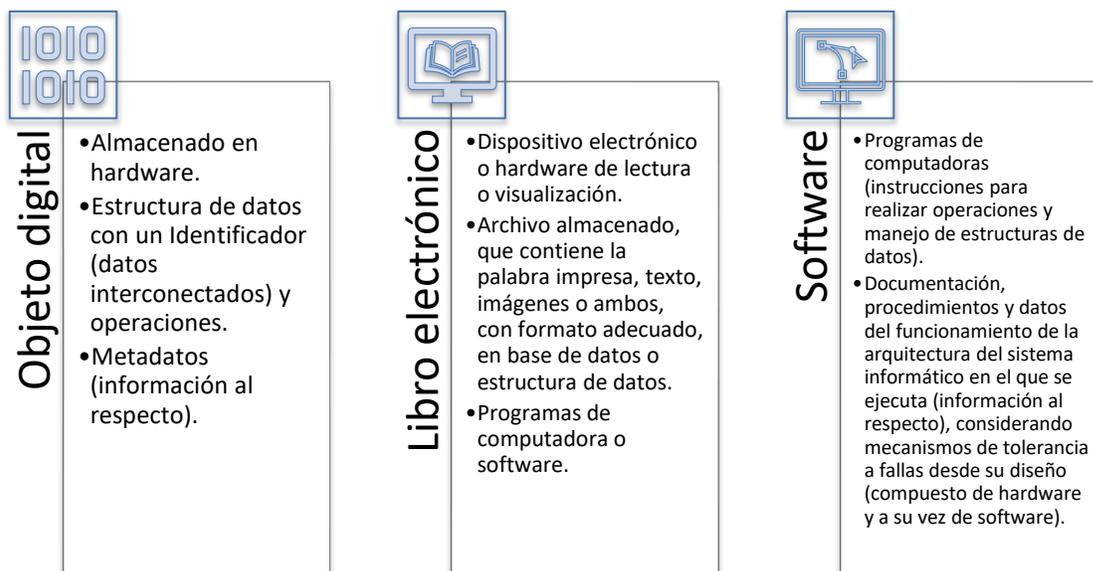
se determina durante su ejecución y el segundo, al momento de su diseño y construcción en donde además se deben contemplar mecanismos de tolerancia a fallas relacionados con su calidad (Wang, 2002b, p. 7). Por lo que es posible identificar dos elementos como componente del software:

1. Programas de computadoras (instrucciones para realizar operaciones y manejo de estructuras de datos)
2. Documentación, procedimientos y datos del funcionamiento de la arquitectura del sistema informático en el que se ejecuta (información al respecto), considerando mecanismos de tolerancia a fallas desde su diseño (compuesto de hardware y a su vez de software)

Esto permite identificar que, en un sistema informático, podemos encontrar muchas piezas de software, y todas estas son programas de computadora que ejecutan operaciones y manejan estructuras de datos, en donde cada pieza debe contar con la documentación de los procedimientos y datos para interactuar con el hardware y más software, del sistema en el que habita.

Con esto, entonces, ahora conviene comparar los elementos identificados al inicio de esta sección referentes al objeto digital, el libro electrónico y el software, Figura 4:

Figura 4. Componentes de objeto digital, libro electrónico y software



Fuente: La autora, enero 2021.

Lo primero que se observa con base en los elementos identificados, es que tanto en el objeto digital como en el software se menciona la necesidad de contener información

referente al elemento y su entorno, sin embargo, se emplean conceptos distintos, metadatos y documentación, por lo que en este momento sólo se reconocerá que, en ambos, es importante contar con información que ayude a entender el entorno en el que se encuentra el objeto o software y más adelante se tratará el tema con profundidad.

Además, el objeto digital vive en un hardware lo cual también es inherente al software y la relación entre estructura de datos y programa de computadora, se establece por que el programa de computadora es el que contiene las instrucciones para operar dicha estructura y este, una vez construido como software se encapsula como un objeto digital.

Respecto al libro digital y el software, se identifica el elemento hardware en ambos, para el libro electrónico, el hardware es el medio que lo hace perceptible, pues se requiere al menos una computadora para poder verlo y además que existan programas u otras piezas de software (sistema operativo y software para lectura) que permiten que este se visualice y lea como una metáfora de libro, pues el archivo por sí sólo, sin el software de lectura y el dispositivo no se puede percibir como libro, de ahí que es posible pensar en que en efecto el libro electrónico se constituye como software diseñado para desenvolverse en un sistema informático y cumplir su misión como abstracción del mundo real.

Esto al compararlo con los datos recabados en la historia del libro electrónico, que se trató en la sección anterior de este capítulo, cobra sentido, pues hay que recordar que los formatos de archivo utilizados para libros (PDF, EPUB, MOBI, etc.) están contruidos con lenguaje de descripción de páginas o PDL, es decir que están codificados en un lenguaje utilizado para indicar a la computadora cómo operar los datos contenidos para construir estructuras o páginas que son parte de un objeto concebido por su creador como libro.

Además, la evolución de los formatos de archivo también muestra cómo estos fueron cambiando para permitir más posibilidades o semejanzas con el libro en papel, y estos formatos de archivo requieren siempre un software lector y mínimo una computadora o sistemas informáticos para los cuales están hechos, con los cuales las personas logramos visualizarlos como libros.

De todo lo anterior, se puede establecer que el libro electrónico o digital desde el punto de vista que en este trabajo de investigación se quiere abordar, se puede definir a partir de los conceptos de Barker, Henke y Wang de la siguiente manera:

*Es un software, objeto digital o tipo especial de información constituido en uno o más formatos de archivo de computadora con comportamiento computacional para abstraer la idea del libro en el mundo material, que requieren de un sistema informático (con nuevas dinámicas en su ciclo de vida y cadena de valor) para lograr interactuar con el mundo físico, es decir, dispositivos de hardware, otras piezas de software (programas de computadora y otros objetos digitales) y documentación asociada como metadatos*

De lo anterior, la necesidad de analizar al libro electrónico o digital desde su naturaleza como software para entender cómo esto afecta su cadena de valor, ya que al tratarse como un producto de software es necesario entender su construcción desde la visión de la Ingeniería de software, pues como Reilly lo menciona:

El software es ahora una industria importante. [...] la necesidad de prestar especial atención a la construcción de software ha dado lugar a la nueva disciplina de ingeniería, la ingeniería de software (Reilly, 2003, p. 238)

De lo anterior, que es necesario identificar cómo la ingeniería de software puede aportar elementos durante la edición del libro que contribuyan a su preservación, principalmente en la construcción de metadatos. Sin embargo, antes de abordar directamente este tema, primero se revisará qué es o qué contempla la preservación de libros electrónicos o digitales.

#### 1.4. LEGISLACIÓN PARA LA PRESERVACIÓN DEL LIBRO DIGITAL

La preservación de acuerdo con la UNESCO<sup>68</sup> es una labor de espectro internacional que permite salvaguardar la memoria en peligro de desaparecer, además forma parte de los impulsos dirigidos por el programa Memoria del Mundo o MoW<sup>69</sup>, que:

tiene por objeto salvaguardar el patrimonio documental mundial, democratizar el acceso a éste y crear una mayor conciencia de su importancia y de la necesidad de preservarlo (UNESCO, 2002, p. 5)

Como ya se mencionó al inicio de este capítulo, el libro, es parte del patrimonio documental de un país porque en sus páginas contiene información respecto a el conocimiento generado por las naciones, de ahí que la defensa y resguardo de este patrimonio regularmente forma parte de las responsabilidades de una biblioteca o archivo nacional como entidad depositaria (UNESCO, 2003, p. 6).

Sin embargo, este trabajo de investigación se encamina sólo al entendimiento de la preservación del patrimonio digital y en este sentido, cabe resaltar que las *Directrices para la Preservación Digital*, publicadas por la UNESCO en 2003, destacan la importancia de modificar o adecuar la legislación para el Depósito Legal en las naciones, considerándolo un elemento esencial como “política nacional de preservación” (UNESCO & BNA, 2003, p. 13) pues esto otorga a las bibliotecas y archivos, mecanismos para garantizar la recepción de las publicaciones digitales, en las que se encuentran incluidas los libros (IPA, 2014).

De ahí que, las palabras que Marianne Scott ex directora de la Biblioteca Nacional de Canadá que a continuación se presentan, ayudan a introducir el punto de partida del tema que aquí se tratará (Larivière, 2000, p. 3):

---

<sup>68</sup> United Nations Educational, Scientific and Cultural Organization

<sup>69</sup> Memory of the World

[...] las bibliotecas nacionales siguen desempeñando un papel esencial en la evolución de la sociedad del conocimiento. Para una biblioteca nacional, encargarse de la adquisición y preservación del patrimonio de publicaciones del país a fin de que todos puedan consultarlo significa enfrentarse con numerosos desafíos. El depósito legal es un importante instrumento que asiste a las bibliotecas nacionales en el cumplimiento de esa función. [...]Sin embargo, la aparición de nuevos formatos, y en particular las publicaciones digitales, ha planteado nuevas cuestiones. Es imprescindible que la información que el público puede consultar en formato digital forme parte de la colección patrimonial de las bibliotecas nacionales. (Larivière, 2000, p. III)

Estas palabras las escribió como prólogo en el documento *Legislación Sobre Depósito Legal: Directrices*, publicado en el año 2000 para referirse a la labor de las bibliotecas nacionales como garantes y depositarias del patrimonio documental y el Depósito Legal como instrumento jurídico que otorga a estas instituciones la posibilidad de preservar a las publicaciones digitales incluyendo a los libros electrónicos o digitales (Larivière, 2000, p. III).

En este trabajo de investigación no se profundizará en las labores de una biblioteca o archivo nacional, por lo que únicamente se menciona de forma general el alcance de estas instituciones y para eso, se toma en consideración lo publicado por la UNESCO desde 1970 en las *Directrices sobre bibliotecas nacionales*<sup>70</sup> en donde, se refiere lo siguiente:

Las bibliotecas que, cualquiera que sea su denominación, son responsables de la adquisición y conservación de ejemplares de todas las publicaciones impresas en el país y que funcionan como bibliotecas “de depósito”, en virtud de disposiciones sobre el depósito legal o de otras disposiciones. Normalmente, pueden desempeñar también algunas de las funciones siguientes: elaborar una bibliografía nacional; reunir una colección amplia y representativa de obras extranjeras, que también comprenda libros relativos al propio país; actuar como centro nacional de información bibliográfica; compilar catálogos colectivos; publicar la bibliografía nacional retrospectiva. [...] (Schick citado por Sylvestre, 1987, p. 1).

De ahí que, una biblioteca nacional o institución depositaria juega un papel importante en la preservación de libros, pues es la responsable de adquirir y conservar dichas obras, a partir de las disposiciones que el Depósito Legal en cada país establezca, por lo que, en este apartado de la investigación, se pretende hablar de forma general de qué es el Depósito Legal y por qué éste requiere cambios para propiciar la preservación del libro electrónico o digital, además de presentar algunos ejemplos de los cambios más relevantes en las legislaciones de países que llevan la delantera en el tema de la preservación digital y esbozar la situación de México en este tema, de tal forma que los datos aquí presentados sirvan para el análisis que se llevará a cabo en el transcurso de esta investigación.

El depósito legal es una obligación establecida por la ley, con arreglo a la cual toda entidad, ya sea esta comercial o pública, y toda persona que produzca cualquier tipo de

---

<sup>70</sup> En este documento además de la UNESCO participó también la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA).

documentación en múltiples ejemplares ha de depositar uno o varios ejemplares en una institución nacional reconocida. (Larivière, 2000, p. 3)

Así pues, de acuerdo con las directrices para la legislación del depósito legal publicado por la UNESCO, el Depósito Legal es un instrumento jurídico que en cada país tiene su propia forma de legislar, sin embargo, éste es un mecanismo por el cual se recaba parte o la totalidad del patrimonio documental de las naciones, pues con éste, en cada región se establece la cantidad de materiales a entregar, así como las características y formatos de entrega e instituciones a las que se entrega, que por lo general son archivos o bibliotecas nacionales (Larivière, 2000, p. 3).

Sin embargo, en el año 2000 la UNESCO con la publicación de la *Legislación Sobre Depósito Legal: Directrices*, recalcó la importancia de que las naciones integrasen consideraciones en sus legislaciones respecto al patrimonio digital, ya que de acuerdo con el mismo documento, el patrimonio digital tiene peculiaridades que requieren un nuevo tratamiento, pues por ejemplo, los libros en línea, son como “un ejemplar único poseído, almacenado y controlado únicamente por los editores/productores” (Larivière, 2000, p. 32) ya que en ese momento, en muchos países no eran entregados a las bibliotecas por lo que el riesgo de perderse o desaparecer, estaba presente si no se lograba resguardar y preservar por las bibliotecas.

Al respecto, en 2014 un reporte generado por el International Publishers Association, IPA, el primer país en adoptar el depósito legal digital de libros electrónicos fue Holanda y fue gracias al trabajo que realizaron de manera conjunta entre la Biblioteca Nacional y la Asociación Holandesa de Editores desde 1999, con lo cual en 2005 lograron los cambios en su legislación y una actualización en 2014. Cabe resaltar que en este caso, se señala en el reporte, que los editores además de entregar los materiales entregan metadatos (IPA, 2014, p. 3).

Por otro lado, una encuesta internacional realizada en 2017 por la Federación Internacional de Asociaciones de Bibliotecarios, IFLA, el Consorcio de Preservación de Internet IIPC, el Instituto de Periodismo Donald W. Reynolds y las Bibliotecas de la Universidad de Missouri, identificaron que “el progreso en la legislación sobre depósito legal en electrónico ha sido lento” (Zarndt et al., 2018, p. 25), pues sólo el 69% recibe el depósito legal digital.

Alemania, es otro ejemplo, que desde el año 2006 hizo cambios en su legislación para incluir a todas las publicaciones electrónicas como parte de su Depósito Legal, el cual es entregado en su Biblioteca Nacional y además en 2009 instauró un enfoque totalmente automatizado para el Depósito Legal, gracias al trabajo con productores y editores, con lo cual, de acuerdo con el mismo reporte, para 2014 les fue posible recopilar entre 1000 y 1500 títulos al día. Como dato interesante, se señala en el reporte del IPA que la ley en Alemania incluye la recolección de material autopublicado y en su sistema automatizado, el editor además del material también envía metadatos (IPA, 2014, p. 2).

Otro ejercicio de revisión en cuanto a las adecuaciones en las legislaciones en depósito legal, es el que publicó en 2018, Biblioteca Jurídica del Congreso de los Estados Unidos, un informe en el que se analizó a 15 países<sup>71</sup>, seleccionados por tener una amplia experiencia con la recolección del Depósito Legal y preservación digital del cual se puede resumir lo siguiente (LLC-EUA, 2018):

- Todos los países tienen leyes que regulan programas obligatorios, voluntarios y pilotos para el depósito y preservación de publicaciones digitales, en línea y nacidas digitales.
- Desde 1999 la mayoría de estos ha desarrollado un programa de depósito electrónico en los últimos diez a quince años.
- En la mayoría de los países encuestados la responsabilidad del depósito recae en los editores o productores y se establecen normas precisas para la entrega de estos materiales en repositorios.
- También en algunos casos el depósito incluye sitios web, publicaciones en línea, publicaciones nacidas digitales y copias digitalizadas de materiales impresos, software y bases de datos.
- Existen casos en el que se permite la sustitución del depósito de una obra impresa por la copia digital.
- Puede haber reglas separadas para películas, materiales sonoros y de gran tamaño.
- En algunos casos, la recolección de sitios web se restringe a URL's que corresponden a los códigos de dominio nacional.
- Hay casos en los que los sitios web privados o de acceso restringido pueden o no incluirse.
- Existen casos en el que se recolectan los libros electrónicos directamente del sitio web del editor.
- El acceso a todo lo depositado depende de la situación comercial y de derechos de autor de la publicación y el responsable del cuidar el acceso y establecer las restricciones necesarias es la institución depositaria.

En contraste con esto, en América Latina, a principios de 2019 el *Centro Regional para el Fomento del Libro en América Latina y el Caribe, CERLALC*, bajo el auspicio de la UNESCO publicó *Recomendaciones para la actualización de las normas sobre el depósito legal en América Latina* <sup>72</sup>(CERLALC-UNESCO, 2019), en las cuales participaron representantes de bibliotecas nacionales de la región y España además de un representante del Grupo Iberoamericano de Editores, GIE, sin embargo, no contiene datos de países que ya tengan avances en este tema, pero el documento expone algunos de los principales problemas

---

<sup>71</sup> Australia, Estonia, Francia, Alemania, Israel, Nueva Zelandia, Corea del Sur, España, Canadá, China, Japón, Reino Unido, Países Bajos y Noruega.

<sup>72</sup> En este documento se menciona que en el año 2017 los libros digitales representaron el 25% de la producción editorial en América Latina

identificados por las instituciones depositarias que de entrada se refiere a la gran cantidad de recursos autopublicados, ya que en 2017 señalan que identificaron un total de 23,474<sup>73</sup> recursos (CERLALC-UNESCO, 2019, p. 12), de ahí que se reitera la recomendación de seleccionar lo que se debe guardar:

El promedio actual de producción de contenidos web por minuto desborda cualquier intento de almacenamiento y catalogación en todos los países de América Latina, por lo que los esfuerzos de conservación patrimonial de páginas y portales web deberán estar inmersos en estrategias de priorización [...] En el estado actual de las capacidades técnica y administrativa, su catalogación y conservación implica una carga que no pueden asumir las entidades depositarias de la región a la luz de las estrategias de priorización y optimización de recursos. Por este motivo, deberá existir una obligación general a cargo de los productores de bases de datos y obras multimedia y transmedia de informar sobre su creación, y realizar el depósito solo en aquellos casos en que lo solicite la entidad competente. (CERLALC-UNESCO, 2019, p. 13).

Respecto a México, el Depósito Legal es recibido por biblioteca del H. Congreso de la Unión<sup>74</sup> y la Biblioteca Nacional<sup>75</sup> y se encuentra legislado por un decreto el cual, hasta el momento de la elaboración de esta investigación, la última modificación que ha tenido fue en 1991 y de acuerdo con Rosa María Fernández de Zamora se hizo con una muy buena visión a futuro, pues:

El decreto de depósito legal vigente en México es del 23 de julio de 1991 [...] es uno de los más avanzados que se conocen pues se previó el depósito de todo tipo de publicaciones electrónicas (Fernández de Zamora, 2000).

Sin embargo, también Fernández de Zamora señaló desde el año 2000 que los cambios tecnológicos han sido tan grandes y el libro electrónico se ha transformado, por lo que es necesario cambiar la legislación (Fernández de Zamora, 2000). De ahí que, tanto la Biblioteca Nacional como el Sistema de Bibliotecas del Congreso de la Unión receptores del Depósito Legal en México han trabajado desde inicios del siglo XXI presentando propuestas para contemplar cambios en el Decreto. La primera propuesta de modificación se presentó en 2002 y además contempló la transformación del decreto en una ley, pues cabe aclarar que, de acuerdo con el Diccionario Jurídico Mexicano, de la Suprema Corte de Justicia de la Nación, el Decreto en México se refiere a:

[...] Toda resolución del Congreso General [...] dentro de la órbita de atribuciones del Poder Legislativo [...] sean sólo respectivas a determinados tiempos, lugares, corporaciones, establecimientos o personas [...] Acto del Poder Ejecutivo referente al modo de aplicación de las leyes en relación con los fines de la administración pública. Disposición de un órgano

<sup>73</sup> De acuerdo con CERLALC esto representa un crecimiento de más de 190% en la última década.

<sup>74</sup> Se tienen registros de ordenamientos jurídicos desde 1846, sin embargo, es hasta 1936 que se establece el Decreto de Depósito Legal emitido por el gobierno de Lázaro Cárdenas, y es hasta 1957 que el decreto establece el depósito tanto a la Biblioteca del Congreso de México como a la Biblioteca Nacional.

<sup>75</sup> La Biblioteca Nacional de México (BNM) se creó en 1867 por el presidente Benito Juárez y forma parte de la Universidad Nacional Autónoma de México (UNAM) desde 1929. En 1944 la UNAM creó formalmente el acervo de la Hemeroteca Nacional de México (HNM) dentro de la Biblioteca y en 1967 la conforma el Instituto de Investigaciones Bibliográficas para encargarse de la administración de ambos acervos documentales, que hasta el momento son responsabilidad de la UNAM.

legislativo que no tiene el carácter general atribuido a las leyes. Resolución judicial que contiene una simple determinación de trámite. (Aguilar y Cuevas citada por Estrada, 2017).

A continuación, se cita el primer argumento de la exposición de motivos presentada ante la Cámara de Diputados en 2002 (Depósito Legal, 2002):

[...] Con el vertiginoso avance tecnológico que en materia de sistemas de almacenamiento y transmisión de información se ha presentado en los últimos años, los documentos de importancia cultural introducen nuevas formas de publicación, si bien el ordenamiento vigente contempla los materiales audiovisuales y electrónicos que contengan información, no se mencionan específicamente materiales que se producen en México, como el software o como los libros electrónicos que sólo están a disposición mediante sistemas como la Internet, razón por la que puede considerarse que el ordenamiento vigente ya no responde a las necesidades actuales del Depósito Legal. (Depósito Legal, 2002)

Esta propuesta principalmente pretendía integrar además de lo considerado ya en el decreto vigente, a los libros electrónicos, bases de datos y software, sin embargo, la propuesta no prosperó y fue turnada a revisiones presentada nuevamente en el 2006 retirando la palabra software y con la modificación en cuanto a los materiales recibidos (Depósito Legal, 2006):

[...] Capítulo II Del Depósito Legal Artículo 3 Las publicaciones que estarán sujetas al depósito legal, enunciativa y no limitativamente, son: a) Libros, tanto de su primera edición como de las siguientes ediciones en sus diferentes presentaciones, siempre que éstas contengan modificaciones respecto de la primera, exceptuándose, en consecuencia, las simples reimpressiones. b) Publicaciones periódicas y seriadas: diarios oficiales, periódicos, revistas, boletines, anuarios y cualquier otra análoga. c) Mapas o planos cartográficos, guías, que contengan especificaciones, señalizaciones o relieves que signifiquen interés para uso legislativo, jurídico, académico, técnico, de investigación o de interés cultural. d) Partituras impresas completas, folletos y otros materiales impresos de contenido cultural, científico y técnico. e) Publicaciones de los Poderes de los tres niveles de gobierno y sus equivalentes en el Distrito Federal, y los estados de la República. f) Microformatos: micropelículas, microfichas. g) Audio, videocasetes, DVDs o cualquier otro tipo de audiograma o videograma realizadas por cualquier procedimiento o sistema empleado en la actualidad o en el futuro. h) Material iconográfico publicado: carteles, tarjetas postales, grabados, fotografías, diapositivas destinadas a la venta, o cualquier otro análogo. i) Disquetes, cintas Dat, DVDs, discos compactos o cintas magnéticas, que contengan información cultural, científica o técnica y/o dispositivos magnéticos, digitales, electrónicos o cualquier otro soporte presente y futuro que registre estos tipos de información, que se edite o grabe con cualquier sistema o modalidad destinado a la venta o que simplemente se publique, y j) Las publicaciones electrónicas, digitales o bases de datos que se hagan públicos por medio de sistemas de transmisión de información a distancia, cuando el origen de la transmisión sea el territorio nacional (Depósito Legal, 2006)

Esta propuesta integraba a las publicaciones electrónicas y bases de datos, sin embargo no prosperó y en 2013 se presentó una nueva propuesta, que también consideraba a las bases de datos a pesar de que sólo intentó derogar el Decreto sin intentar crear una ley, el principal cambio fue que pretendía que los editores hicieran el depósito legal en papel y digital de todas sus publicaciones, aquí se presenta una parte del texto presentado (Decreto de Depósito Legal, 2013):

En el decreto se dispone la obligación de los editores de hacer depósito legal digitalizado de sus publicaciones. [...] contempla los materiales que estén disponibles al público mediante sistemas de transmisión de información digital o cualquier otro medio y sea producida en el país. Deben entregar ejemplares en versión digital e impresa de cada publicación de las ediciones y producciones de sus obras, a la Biblioteca Nacional de México, a la Hemeroteca Nacional de México, al Sistema de Bibliotecas del honorable Congreso de la Unión, y a la Biblioteca del Congreso de la entidad federativa en la que tenga su sede legal la casa editora. Un ejemplar de todas las publicaciones electrónicas, digitales o bases de datos que se hagan públicos por medio de sistemas de transmisión de información a distancia, cuando el origen de la transmisión sea el territorio nacional. Las bibliotecas depositarias podrán solicitar el depósito de publicaciones con valor crítico para el patrimonio cultural y editorial de la nación (De Decreto de Depósito Legal, 2013).

Esta propuesta a diferencia de la de 2002 contemplaba como obligación la entrega de todas las publicaciones en formato impreso y digital, esto desde luego que implicaba un cambio significativo ya que el formato digital no era alternativo. Algo interesante de señalar es la reacción que provocó en los editores, de tal manera que el presidente de la Cámara Nacional de la Industria Editorial Mexicana o CANIEM, José Ignacio Echeverría consideró ilegal la propuesta argumentando que:

[...] los editores no siempre tienen la autorización de los autores para realizar la edición digital, y con la posibilidad de que se pongan en línea materiales considerados de interés científico o cultural se haría a los editores perder el control de las obras que editan y publican (Jaubert, 2014).

A pesar de lo anterior, el 30 de noviembre de 2017 la Cámara de Diputados de la LXIII Legislatura, presentó una propuesta que buscó una modificación al decreto y contempló las siguientes adiciones con respecto a los recursos electrónicos (Reforma de Decreto de Depósito Legal, 2017):

Artículo Segundo. [...] Esta obligación se cumple con la entrega de ejemplares de cada una de las ediciones y producciones de sus obras, a la Biblioteca Nacional y a la Biblioteca del Congreso de la Unión, en los términos señalados en el artículo tercero del presente decreto. Así como la versión electrónica depositada en los acervos digitales de dichas bibliotecas. [...] Artículo Tercero. [...] C) Para el caso de la entrega vía web, es necesario contar con un registro que permita tener un orden temporal de las entregas de dichos materiales, mismos que podrán ser considerados de manera automática para consulta ciudadana. [...] Artículo Quinto. La Biblioteca Nacional y la Biblioteca del Congreso de la Unión, deberán: F) Apoyar los derechos de autor asegurando que el uso que se haga de dichos materiales no violenta los derechos de autor de cualquiera que sea su aportación al acervo cultural nacional. [...] Artículo Decimotercero. Las personas que deseen consultar la versión electrónica de las obras bajo Depósito Legal deberán cubrir una cuota destinada al pago de los derechos de autor, así como al mantenimiento de la plataforma electrónica para su consulta. El pago de dicha cuota no implicará la descarga de la obra, sino su consulta en línea, sin derecho a impresión, por un periodo de 14 días hábiles. Artículo Decimocuarto. La cuota para la consulta electrónica será ajustada dependiendo de la cantidad de recursos electrónicos que el usuario desee utilizar, mientras que la consulta física dentro de las Bibliotecas será gratuita y controlada (Reforma de Decreto de Depósito Legal, 2017).

Esta propuesta tampoco prosperó, sin embargo, la Biblioteca Nacional y el Sistema de Bibliotecas del Congreso de la Unión, realizaron actividades de diálogo para conciliar los diferentes intereses en la búsqueda de acuerdos y trabajaron en nuevas propuestas conformando grupos de expertos de ambas y otras instituciones (Galina Russell et al., 2019), para crear una nueva propuesta de ley, sin embargo, estos esfuerzos no se vieron cristalizados, pues la Biblioteca Nacional de México publicó en el informe de labores de 2018 lo siguiente:

Para fortalecer la recepción de la producción editorial mexicana, analógica y digital, en 2018 la Biblioteca y la Hemeroteca, junto con el Instituto, celebraron reuniones de trabajo con las instituciones involucradas (Dirección General de Servicios de Documentación, Información y Análisis de la Cámara de Diputados, la Fonoteca Nacional, la Filmoteca de la UNAM y el Fideicomiso de la Cineteca Nacional), con la finalidad de dar continuidad al proyecto de Ley de Depósito Legal en el cual se ha venido trabajando desde 2016. Se concluyó el borrador de la propuesta de Ley y se avanzó en un glosario de términos y en una exposición de motivos. Sin embargo, por circunstancias políticas nacionales y universitarias, el proyecto mencionado no se presentó al Legislativo (BNM, 2019, p. 13).

Sin embargo, en 2018 se presentó una propuesta derivada de lo planteado en 2013, que tampoco prosperó, pero además no integraba las consideraciones que trabajó la Biblioteca Nacional en colaboración con otras instituciones y nuevamente sólo contempló la actualización del decreto, manteniendo la intención de recopilar versiones impresas y digitales de todos los editores (Comisión de Cultura, 2018).

[...] Artículo Segundo. De los efectos del artículo anterior todos los editores y productores de materiales bibliográficos y documentales, están obligados a contribuir a la integración del patrimonio editorial de la Nación. Esta obligación se cumple con la entrega de ejemplares en versión digital e impresa de cada publicación de las ediciones y producciones de sus obras, a la Biblioteca Nacional de México, al Sistema de Bibliotecas del H. Congreso de la Unión y a la Biblioteca del Congreso de la Entidad Federativa en la que tenga su sede legal la casa editora, en los términos señalados en el Artículo Tercero del presente Decreto [...] (Comisión de Cultura, 2018).

Finalmente, de acuerdo a lo publicado en el sitio web del canal del congreso en México<sup>76</sup> en año 2021 se aprobó reformar la Ley General de Bibliotecas, con la cual se abroga el Decreto de Depósito Legal de 1991 e integra en el articulado de dicha ley, lo correspondiente al Depósito Legal (Congreso de la Unión México, 2021):

[...] En sesión semipresencial, la Cámara de Diputados aprobó el dictamen a la minuta que expide la Ley General de Bibliotecas, la cual tiene como objetivo establecer las bases de coordinación de los gobiernos Federal, de las entidades federativas, los municipios y alcaldías de la Ciudad de México en materia de bibliotecas públicas [...] La nueva ley también tiene como objetivo definir las políticas de establecimiento, sostenimiento y organización de las bibliotecas públicas [...] También, fomentar y garantizar la conservación del patrimonio documental, bibliográfico, hemerográfico, auditivo, visual, audiovisual, digital y, en general, cualquier otro medio que contenga información afín, estableciendo instrumentos para la difusión cultural, la consolidación de la memoria comunitaria y el

76

[https://www.canaldelcongreso.gob.mx/noticias/14221/Diputados\\_avalan\\_expedir\\_nueva\\_Ley\\_General\\_de\\_Bibliotecas](https://www.canaldelcongreso.gob.mx/noticias/14221/Diputados_avalan_expedir_nueva_Ley_General_de_Bibliotecas)

progreso educativo y regular los términos del Depósito Legal [...] El dictamen aprobado con 443 votos a favor, uno en contra y cero abstenciones, se remitió al Ejecutivo federal para sus efectos constitucionales (Congreso de la Unión México, 2021).

Algunas de las modificaciones presentadas en esta propuesta de Ley General de Bibliotecas son (Ley General de Bibliotecas, 2020):

Artículo 2. Para los efectos de la presente Ley, se entenderá por: [...] XIV. Libro: Toda publicación unitaria, no periódica, de carácter literario, artístico, científico, técnico, educativo, informativo o recreativo, editada en cualquier soporte, lenguaje o código, incluido el digital, cuya edición se haga en su totalidad de una sola vez en un volumen o a intervalos en varios volúmenes o fascículos. Comprenderá también los materiales complementarios en cualquier tipo de soporte, incluido el electrónico, que conformen, conjuntamente con el libro, un todo unitario que no pueda comercializarse separadamente. [...] XV. Preservación: Conjunto de actividades administrativas y económicas orientadas a prevenir el deterioro de los documentos, garantizando así la permanencia física de los acervos y de la información contenida en ellos. [...] XVI. Publicación electrónica: Toda información unitaria, no periódica, de carácter literario, artístico, científico, técnico, educativo, informativo o recreativo, editada en forma digital o existente en versión electrónica, fijada o no en algún soporte, cuyo contenido es susceptible de contar con una edición impresa. Por sus características presenta modalidades de consulta y acceso sujetas a la licencia de uso, temporalidad y dispositivos de lectura disponibles de cada biblioteca. [...] CAPÍTULO X Del Depósito Legal de Publicaciones [...] Artículo 33. Se declara de interés público la recopilación, integración, almacenamiento, custodia y conservación de toda obra de contenido educativo, cultural, científico, técnico o de esparcimiento, distribuida para su comercialización o de manera gratuita, en formatos impreso o electrónico, analógico o digital, en el territorio nacional. El conjunto de obras recopiladas constituye el Depósito Legal. Artículo 34. Las obras a que se refiere el artículo anterior podrán ser, de manera enunciativa y no limitativa, las siguientes: I. Libros, publicaciones periódicas, catálogos, folletos y pliegos; II. Publicaciones periódicas como periódicos, diarios, anuarios, revistas, y memorias; III. Material cartográfico como mapas y planos, cartas de navegación, aeronáuticas o celestes; IV. Partituras; V. Fonogramas, discos y cintas; VI. Obras audiovisuales, micropelículas, diapositivas y fotografías; VII. Material gráfico, carteles y diagramas, y VIII. Cualquier otra que se considere relevante para documentar la memoria del conocimiento en el territorio nacional. [...] (Ley General de Bibliotecas, 2020).

Cabe señalar que esta última propuesta establece definiciones para libro y publicación electrónica que no existían en las propuestas anteriores, sin embargo la definición de libro incluye al digital pero lo separa del concepto de publicación electrónica, ya que ésta última de acuerdo a la ley “presenta modalidades de consulta y acceso sujetas a la licencia de uso, temporalidad y dispositivos de lectura disponibles de cada biblioteca” (Ley General de Bibliotecas, 2020), por lo que resulta confuso y no queda claro si el libro entra o no en esta definición. Además, lo más importante, es que la publicación electrónica, no se contempla como concepto, en el capitulo del sujeto de depósito legal del artículo 34 de dicha ley<sup>77</sup>.

<sup>77</sup> Cabe recalcar que al revisar la nueva Ley General de Bibliotecas esta no contiene elementos que coadyuven a la protección del patrimonio digital pues no refiere nada respecto a la entrega de metadatos, como sucede en otros países, de hecho, el definir por separado a la publicación electrónica del concepto de libro y no utilizar este término dentro del capitulo del Depósito Legal y prácticamente en ningún lugar de la ley, parece que lo excluye de la misma, pues se usa el termino digital y electrónico como

En este trabajo de investigación no se pretende hacer un análisis profundo del Depósito Legal Digital en México, pues no es el objetivo, sin embargo, como se verá en los siguientes capítulos, es un factor importante para llevar a cabo la preservación y por lo tanto, para el modelo que aquí se expone, pues la entidad receptora de Depósito Legal requiere una figura jurídica que le permita impactar en la edición del libro digital y en los metadatos que producen los editores, como sucede en los países que han logrado avances en el tema y que ya se mencionaron en el desarrollo de esta sección.

Sin embargo, este tema se retomará más adelante, por lo pronto, ahora es necesario hablar de la preservación digital del libro para entender qué es, qué implica y cómo impacta ésta, en las organizaciones y en las naciones que deciden llevarla a cabo.

### 1.5. ORIGEN DE LA PRESERVACIÓN DIGITAL

Cabe señalar que el concepto de preservación digital se ha transformado con el paso del tiempo, por lo que en esta sección se tratará de dar un esbozo general de su origen y metamorfosis, además de analizar cuál es el alcance que tiene en las organizaciones.

La preservación digital surgió como el proceso de reproducción o conversión del libro impreso a través de la digitalización a principios de los años 90's<sup>78</sup> y tuvo un gran impulso principalmente en Estados Unidos y la Unión Europea, de ahí que, hay quienes la definen como:

[...] la técnica de digitalización de documentos en soporte papel para asegurar su conservación en el tiempo y evitar el deterioro del original [...] (REBIUN, 2009, p. 4)

La preservación digital es la exploración de enfoques sistemáticos para ingerir, archivar y disseminar información digitalizada (Luan et al., 2010, p. 1)

Desde esta perspectiva, la digitalización se consideró un medio de preservación que garantizaba la permanencia del material original impreso ya que, desde dicho punto de vista, reduce la manipulación y permite un mayor acceso a través de la reproducción en digital. De hecho, se pensó que la digitalización se convertiría en un soporte con mayor permanencia que el microfilm, de ahí que, por ejemplo, la Federación de Bibliotecas Digitales<sup>79</sup> en Estados Unidos, al conformarse como Consorcio de Preservación Digital declaró que:

---

calificativo del libro y la publicación, pero en ningún momento se emplea como el concepto que se define al inicio de la ley a pesar de que incluyen la entrega de versiones digitales. Además, respecto a la figura de la Biblioteca Nacional, esta ley no le atribuye o reconoce como entidad con las posibilidades de establecer criterios de entrega de los materiales, sólo le otorga la responsabilidad de organizar y poner a disposición su propia recepción. Tampoco establece la responsabilidad del gobierno federal para otorgar recursos económicos a la Biblioteca Nacional para llevar a cabo sus labores de preservación en general y digital.

<sup>78</sup> Consorcio de Preservación Digital (<http://cool.conservation-us.org/byauth/waters/dpcmiss.html>) que ahora se convirtió en Digital Libraries Federation o la Federación de Bibliotecas Digitales (<https://www.diglib.org/>). Es una agrupación de universidades norteamericanas encaminadas a impulsar el uso de la tecnología digital para preservar y facilitar el acceso a obras impresas de importancia nacional e internacional en los Estados Unidos de Norte América.

<sup>79</sup> *Digital Preservation Consortium* (<http://cool.conservation-us.org/byauth/waters/dpcmiss.html>) agrupación de universidades norteamericanas encaminadas a impulsar el uso de la tecnología digital para preservar y facilitar el acceso a obras impresas de

La acumulación de evidencia sugiere que la tecnología digital puede estar emergiendo como una herramienta efectiva y asequible de preservación y acceso. Si la evidencia es válida, las instituciones nacionales de investigación necesitarán incorporar esta nueva herramienta [...] <sup>80</sup>(Waters et al., 1994)

Por lo que la digitalización o reproducción en digital ha sido impulsada fuertemente en el mundo, en muchas bibliotecas y centros de documentación incluida América Latina. Por lo que, ésta y la producción de libros electrónicos nacidos en digital, han acelerado la acumulación de contenidos digitales tanto en la web como en medios magnéticos y desde el año 2000, como ya se vio, la cantidad de formatos y posibilidades de construir contenidos digitales comenzó a diversificarse.

Sin embargo, se identificó que el volumen de datos en Internet, además de crecer constantemente, también es difícil de conservar, pues en el año 2000 por ejemplo, se estimó que la información en internet tenía un promedio de vida muy corta, calculada entre 44 días y dos años en el mejor de los casos (UNESCO & BNA, 2003, p. 3), por lo que esto y la obsolescencia de los medios de almacenamiento junto con los programas informáticos, evidenciaron la vulnerabilidad de los objetos digitales y de ahí, que su permanencia se convirtió en una preocupación por lo que, en el año 2003, la UNESCO hizo un llamado para atender y proteger un nuevo tipo de patrimonio, el digital, en un documento denominado *Carta para la Preservación del Patrimonio Digital*:

El patrimonio digital del mundo corre el peligro de perderse para la posteridad. Contribuyen a ello, entre otros factores, la rápida obsolescencia de los equipos y programas informáticos que le dan vida, las incertidumbres existentes en torno a los recursos, la responsabilidad y los métodos para su mantenimiento y conservación y la falta de legislación que ampare estos procesos (UNESCO, 2003, pp. 2, anexo I)

Además, junto con esto la UNESCO publicó las *Directrices para la preservación del patrimonio digital*<sup>81</sup>, en el que se declaró:

preservación digital designa los procesos a que se recurre con objeto de conservar información y cualquier otro tipo de patrimonio existente en forma digital [...] (UNESCO & BNA, 2003, p. 21)

Con lo cual, además de hacer un llamado para atender al patrimonio digital se cambió la concepción del término preservación digital, indicando lo siguiente:

[...] este término no se refiere al uso de imágenes digitales o de técnicas de captura para hacer copias de elementos no digitales, inclusive si éstas se realizan con fines de preservación. Obviamente, la realización de copias digitales (también denominada

---

importancia nacional e internacional en los Estados Unidos de Norte América, esta organización se convirtió en lo que ahora es *Digital Libraries Federation* o la Federación de Bibliotecas Digitales (<https://www.diglib.org/>).

<sup>80</sup> Traducido por Ana Yuri Ramírez

<sup>81</sup> El 19 de agosto de 2003 en París en la 32a reunión, ante 47 estados miembro y 15 organizaciones internacionales, Colin Webb de la Biblioteca Nacional de Australia presentó las Directrices para la Preservación del Patrimonio Digital.

digitalización) puede perfectamente producir materiales de patrimonio digital que necesiten ser preservados[...] (UNESCO & BNA, 2003, p. 21)

De ahí que, a partir de ese momento las digitalizaciones generadas por las instituciones se convirtieron en el objetivo de la preservación, dejando de ser un medio de preservación, pues se comprobó que el medio digital no asegura la permanencia de la información contenida en él y digitalizar, ya no significa preservar digitalmente, de tal forma que ahora las imágenes o archivos de computadora de los libros generados por la digitalización son parte del patrimonio digital declarado por la UNESCO<sup>82</sup> desde 2003 y requieren ser atendidas para permanecer a largo plazo, a partir de lo establecido en las *Directrices para la preservación del patrimonio digital*<sup>83</sup>, que define a la preservación digital como:

[...] el conjunto de los procesos destinados a garantizarla continuidad de los elementos del patrimonio digital durante todo el tiempo que se consideren necesarios (UNESCO & BNA, 2003, p. 37)

[...] la preservación digital consiste en los procesos destinados a garantizar la accesibilidad permanente de los objetos digitales [...] (UNESCO & BNA, 2003, p. 37)

Derivado de las directrices de la UNESCO, diversos países y especialistas que desde ese momento y hasta ahora, siguen explorando las implicaciones y complicaciones de la preservación de los objetos digitales, han desarrollado conceptos para explicar el alcance de la preservación digital, algunos ejemplos son:

El proceso específico para mantener los materiales digitales durante y a través de las diferentes generaciones de la tecnología a través del tiempo, con independencia de donde residan<sup>84</sup> (InterPARES 3, 2012)

[...] es la gestión permanente del contenido digital para garantizar su acceso continuo a lo largo del tiempo. (Library of Congress, 2016)

La preservación digital es el conjunto de principios, políticas, normas y estrategias diseñadas para asegurar que un objeto digital permanezca accesible, inteligible y usable a través del tiempo y de los cambios tecnológicos, su fiabilidad y exactitud están protegidas y su autenticidad es verificable. (InterPARES, 2017, p. 17 traducida por Barnard, Delgado y Voutssás)<sup>85</sup>

La primera definición elaborada por InterPARES es muy parecida a la establecida por la UNESCO, sin embargo, esta emplea el término mantenimiento, que se refiere a los procesos o estrategias técnicas empleadas para modificar un componente de software y adaptarlo a

---

<sup>82</sup> *United Nations Educational, Scientific and Cultural Organization*

<sup>83</sup> El 19 de agosto de 2003 en París en la 32a reunión, ante 47 estados miembro y 15 organizaciones internacionales, Colin Webb de la Biblioteca Nacional de Australia presentó las Directrices para la Preservación del Patrimonio Digital.

<sup>84</sup> Multilingual Archival Terminology, <http://www.ciscra.org/>

<sup>85</sup> Traducción hecha por los autores de un documento de InterPARES The International Research on Permanent Authentic Records in Electronic Systems.

un entorno cambiado<sup>86</sup> (SEVOCAB et al., 2021c), por lo tanto, en esta definición los procesos a los que se refiere implican el hacer cambios sobre el objeto para dar continuidad al mismo.

El segundo concepto, de la biblioteca del Congreso de los Estados Unidos, en comparación con lo establecido por la UNESCO habla de gestión permanente, en lugar de procesos, esto parecería similar, sin embargo, la gestión implica una visión organizacional, administrativa y de control (Chávez, 2018), a diferencia de los procesos o procedimientos, que sólo “establecen el orden cronológico y la secuencia de actividades que deben seguirse para realizar un trabajo” (Münch, 2005, p. 100), que en efecto obedecen y se ejecutan dentro de una visión organizacional, sin embargo podrían verse de forma aislada y reducidos sólo a la parte operativa o técnica, como por ejemplo hacer copias y respaldos de la información.

De ahí que, la definición de la Biblioteca del Congreso de los Estados Unidos contempla una visión más amplia y esto concuerda también con la última definición, establecida por interPARES y traducida por Barnard, Delgado y Voutssás, no obstante, ésta última agrega más elementos, pues habla de principios, normas y políticas, que obedecen al conjunto de elementos que rigen y dirigen el comportamiento de una organización o persona, de tal forma que esta definición además incluye elementos externos a la organización y emplea el término “estrategia”, que no sólo se refiere a la gestión, pues de acuerdo con Henry Mintzberg, especialista en proceso estratégico, la estrategia es una proyección a futuro que implica una planeación y seguimiento:

Una estrategia es el patrón o plan que integra las principales metas y políticas de una organización y, a la vez, establece la secuencia coherente de las acciones a realizar [...] ayuda a poner orden y asignar con base en atributos y deficiencias, los recursos de una organización con el fin de lograr una situación viable y original, así como anticipar los posibles cambios en el entorno [...](Mintzberg et al., 1997, p. 7)

De lo anterior, que la preservación digital no se reduce a procesos sólo operativos de respaldo o copia, transformación de objetos y acceso o consulta, además implica una afectación o cambio organizacional, de ahí que Juan Voutssás, especialista mexicano en preservación digital señala que para establecer procesos de preservación es necesario analizar el medio en el que los objetos digitales son creados, distribuidos y resguardados y entonces adaptar los estándares, recomendaciones y modelos establecidos para el cuidado de los materiales digitales a las organizaciones con base en sus necesidades y posibilidades económicas, geográficas, jurídicas, políticas, técnicas y sociales (Voutssás, 2011).

Esto concuerda con lo visto anteriormente, pues será necesario que cada país u organización revise y adapte las legislaciones que correspondan (por ejemplo derechos de autor y de Depósito Legal cuando se trata de entidades depositarias), las políticas (públicas o de una organización como responsabilidades y/o compromisos), presupuestos y desde luego, las políticas que se refieren a la producción o construcción de objetos digitales o

<sup>86</sup> Del estándar ISO/IEC25051:2014 de Ingeniería de software - Requisitos de calidad y evaluación de productos de software.

concretamente libros digitales, para contemplar la preservación a largo plazo (CCSDS, 2012, pp. 1-3).

El último elemento importante respecto a las definiciones analizadas se relaciona con los riesgos identificados en este tipo de objetos, la pérdida de autenticidad, fiabilidad y exactitud derivados de los cambios en la tecnología y que establecen como metas de la preservación, conservar las características de acceso, uso e inteligibilidad de los objetos a partir de aplicar procesos o estrategias que permitan mantener dichas características. Las directrices de la UNESCO en el apartado *Comprender la preservación digital*, lo menciona de la siguiente manera:

[...] es necesario encontrar las maneras de representar lo que se había presentado originalmente a los usuarios mediante un conjunto de equipos y programas informáticos que permiten procesar los datos. Para lograrlo, es necesario que la comprensión y la gestión de los objetos digitales se realice considerándolos desde cuatro puntos de vista: como fenómenos físicos, como codificaciones lógicas, como objetos conceptuales comprensibles para el ser humano y como conjuntos de elementos esenciales que deben ser preservados para ofrecer a los futuros usuarios lo esencial del objeto (UNESCO & BNA, 2003, p. 37)

Para el combate de estos riesgos se han aplicado estrategias tecnológicas o prácticas que ayudan a la prevención, control y contención de algunas de estas vulnerabilidades, que van desde los respaldos de información, renovación de los soportes de almacenamiento, replicación, respaldo en analógico, utilización de medios persistentes, arqueología digital para el rescate de medios magnéticos; migración de versiones, datos y plataformas; normalización y estandarización; emulación y encapsulamiento, virtualización, certificación o aplicación de sellos, firmas digitales o sumas de verificación digital y la generación de metadatos (Rivera D., 2009, pp. 20-41).

Cabe resaltar que cada una de estas estrategias tecnológicas sigue procesos muy concretos y son aplicables de acuerdo con el tipo y características de cada objeto digital, por lo que en su mayoría deben establecerse con base en una planeación y análisis de los recursos y capacidades de cada institución y las problemáticas o riesgos identificados en el objeto.

De todo lo anterior, que, para fines de este trabajo de investigación:

*La preservación digital es una estrategia organizacional que implica orientar principios, políticas y normas (económicas, legales, organizacionales, tecnológicas y de producción) dentro de las instituciones para gestionar la permanencia de un objeto digital (que incluye a los libros electrónicos o digitales y digitalizados de un impreso, entre otros) contemplando procesos y técnicas estructuradas y perfeccionadas a través del tiempo y los avances tecnológicos configuradas como estrategias tecnológicas para el mantenimiento de los objetos, encaminadas al cuidado de la autenticidad, fiabilidad y exactitud de éstos para lograr conservar en lo posible, las características de acceso, uso e inteligibilidad original (manera en que fue concebido para su acceso, percepción y utilización).*

A continuación, se dará un esbozo general de los conceptos, estándares, modelos y estrategias que en el ámbito de la preservación digital son aplicables dentro de las organizaciones para establecer sus procesos de preservación digital que servirán como marco de referencia para este trabajo de investigación.

## 1.6. LA PRESERVACIÓN DIGITAL Y EL LIBRO DIGITAL

En esta sección se decidió primero ubicar a las organizaciones y trabajos más sobresalientes en el tema de preservación digital, esto para identificar y tomar en consideración los avances y recomendaciones, que dichos organismos hacen respecto de la preservación del libro. Y al final, se hace una descripción del modelo OAIS y otros modelos que son fundamentales para entender la preservación digital del libro.

### 1.6.1. ORGANIZACIONES Y TRABAJOS SOBRESALIENTES EN PRESERVACIÓN DIGITAL

En Europa surgió una de iniciativas más antiguas en el tema, conocido como DLM Forum, el cual hasta el momento se conforma por representantes de la administración pública y de los servicios de archivos nacionales, así como representantes de la industria y la investigación en la unión europea. Ellos son los creadores del estándar Moreq<sup>87</sup> e impulsaron el proyecto E-ARK<sup>88</sup>. Además, en Europa existen diversas organizaciones y grupos trabajando en el tema, por ejemplo, la DPE<sup>89</sup> o Digital Preservation Europe, que ya no existe como tal, sin embargo, fue creada por la Comisión Europea con la participación de instituciones culturales líderes de los Países Bajos, Italia, Alemania, Escocia, la República Checa, Austria, Lituania y Dinamarca.

La DPE dejó de recibir recursos en 2009 y pasó a manos del HATII<sup>90</sup> o Instituto de Tecnología e Información Avanzada en Humanidades de la Universidad de Glasgow, el cual tampoco existe como tal, pues en 2017 se convirtió en Digital Media & Information Studies y actualmente ofrece, programas de investigación en Tecnologías de la Información, Humanidades Digitales, Preservación Digital, Gestión de Archivos y Ciencia Forense Digital. Sin embargo, uno de los trabajos señalados por la DPE como fundamental, fue el fomentar la conciencia pública acerca de la preservación digital en sus comunidades y así cautivar a los investigadores generándoles interés para invitarlos a proponer ideas que resolvieran problemas reales de preservación.

De ahí que, la DPE invirtió muchos recursos en armar productos como videos dirigidos al público en general, primero para atraer al público en el tema y además generar el interés

<sup>87</sup> Modular Requirements for records systems. Norma que establece requisitos en un software para la gestión de documentos. <https://www.moreq.info/>

<sup>88</sup> European Archival Records and Knowledge Preservation. Se centró en identificar los problemas asociados con las tecnologías, sistemas y prácticas independientes de mantenimiento de registros, por lo cual estableció criterios para el apoyo de la preservación digital. <https://www.eark-project.com/>

<sup>89</sup> Digital Preservation Europe, que fue un conjunto de instituciones académicas y culturales europeas que se concentraron en generar conocimientos y recursos de preservación digital para compartirlos.

<sup>90</sup> The Humanities Advanced Technology and Information Institute. fundado en 1997, sin embargo ya no existe y en su lugar ahora existe el DMIS <https://www.gla.ac.uk/undergraduate/degrees/digitalmedia/>

de sus investigadores, para lo cual también organizó eventos para premiar a los investigadores que proponían las mejores soluciones a los problemas planteados. Como se dijo, la DPE ya no está activa, pero sin duda su trabajo debió generar los resultados esperados pues actualmente en Europa existen varias organizaciones que siguen trabajando como la OPF<sup>91</sup> y el DPC<sup>92</sup>.

La DPC, por ejemplo, creada en el Reino Unido, ha logrado integrar a más países incluso de América y Asia, generando una comunidad internacional más amplia, por lo que incluso ha logrado impulsar, un día al año, la celebración mundial de la preservación digital.

Otra iniciativa europea se conformó en 2011 en Estonia, para formar la ANADP<sup>93</sup>, que tuvo por objetivo crear mecanismos para la cooperación entre sus participantes, una de sus primeras acciones fue generar un registro internacional de herramientas y servicios conocido como COPTR<sup>94</sup> y también utilizado por DPC, para establecer un marco de trabajo común y crear un catálogo de posibilidades de capacitación y educación.

Además, la ANADP trató de establecer un modelo distribuido de preservación digital entre los participantes, basado en OAIS<sup>95</sup>, recopilando datos que permitieran cuantificar los costos de inversión que requiere la Preservación Digital y así, integrar instituciones interesadas en financiar proyectos transnacionales, para unir esfuerzos y preservar los recursos de forma cooperativa, sin embargo, durante el desarrollo de esta investigación no se encontró la continuidad del proyecto.

Otra de las iniciativas con mucho tiempo y experiencia en el tema, es InterPARES<sup>96</sup>, especializada en la preservación de archivos administrativos conformada en EUA, ha impulsado la integración de equipos de trabajo<sup>97</sup> en varias naciones, incluyendo América Latina y el Caribe, para fomentar la creación y divulgación de productos académicos originales de libre expresión e innovadores, principalmente para el ámbito de los documentos de archivo y también aplicables en lo general (Voutssás, 2014). Es importante, además, resaltar el trabajo realizado por el grupo InterPARES México y su participación en la ley de archivos y datos personales, pues contemplar aspectos que coadyuvan en la Preservación Digital de los archivos, en el país.

Otro grupo muy importante a nivel mundial, pero creado en Norte América, es el NDIIPP ahora conocido como NDSA que ha impulsado los niveles LoP<sup>98</sup> (NDSA - DLF, 2019) de aplicación general para evaluar programas de preservación digital. Además en Estados Unidos, también está la IDPF ahora PUBLISHING@W3C (W3C, 2021), responsable de crear

---

<sup>91</sup> Open Preservation Foundation

<sup>92</sup> Digital Preservation Coalition

<sup>93</sup> The Aligning National Approaches to Digital Preservation

<sup>94</sup> Community Owned digital Preservation Tool Registry (COPTR) [http://coptr.digipres.org/Main\\_Page](http://coptr.digipres.org/Main_Page)

<sup>95</sup> OAIS modelo para Preservación Digital

<sup>96</sup> The International Research on Permanent Authentic Records in Electronic Systems (InterPARES)

<sup>97</sup> InterPARES Trust (ITrust)

<sup>98</sup> The Levels of Digital Preservation (LoP). Recurso para los profesionales de la preservación digital cuando crean o evalúan su programa de preservación digital.

el formato EPUB como mecanismo para, “compartir publicaciones, guardarlas” y “conservarlas”(W3C, 2021).

De todo lo anterior que, para fines de esta investigación, se presenta una tabla elaborada como una muestra de la cantidad de organismos u organizaciones que se identificaron con una participación en la preservación digital. En esta tabla, se presentan en orden cronológico de acuerdo con el año de creación, acompañado del nombre y siglas, el sitio web, la región o lugar de origen, una descripción tomada de sus propias páginas web, así como la orientación, intereses o campo de estudio, además de señalar si aún siguen activas, ver Tabla 1.

Tabla 1. Organizaciones interesadas en la preservación digital a nivel mundial, identificadas durante el desarrollo de esta investigación.						
Año de creación	Nombre	Lugar de origen	Sitio Web	Descripción tomada del sitio web	Tipo de objeto que buscan preservar	¿Sigue activa?
1996	DLM Forum (Document Lifecycle Management)	Comunidad Europea	<a href="http://www.dlmforum.eu/">http://www.dlmforum.eu/</a>	Foro multidisciplinario que busca resolver los problemas de gestión, almacenamiento, conservación y recuperación de datos legibles por máquina. Está compuesto por representantes de la administración pública y de los servicios de archivos nacionales, así como representantes de la industria y la investigación.	Documentos de Archivo	SI
1999	DLF (Digital Library Federation)	EUA	<a href="https://www.diglib.org/">https://www.diglib.org/</a>	Comunidad que promueve la investigación y aprendizaje a través del diseño de bibliotecas digitales.	Bibliotecas Digitales	SI
1999	InterPARES (The International Research on Permanent Authentic Records in Electronic Systems)	UK, EUA	<a href="http://www.interpares.org/">http://www.interpares.org/</a>	El Proyecto InterPARES 1 se basó en el Proyecto UBC. Fue realizado por investigadores de la Universidad de Columbia británica de 1994 a 1997, en colaboración con el Departamento de Defensa de los Estados Unidos, y tuvo como objetivo establecer estándares para crear registros electrónicos confiables y mantener su autenticidad durante su vida activa y semiactiva.	Documentos de Archivo	SI
1999 a 2016	IDPF (International Digital Publishing Forum)	EUA	<a href="http://idpf.org/">http://idpf.org/</a>	Asociación mundial de comercio y normas para la publicación electrónica. La misión de IDPF era fomentar la adopción global de un ecosistema de publicación digital abierto, accesible e interoperable que permite la innovación. La estrategia de IDPF para avanzar de manera efectiva en su misión fue desarrollar (en colaboración) y promover EPUB como el ecosistema universal de intercambio y entrega accesible para libros electrónicos y otras publicaciones digitales que promueven la plataforma Web Abierta.	Publicaciones electrónicas	NO
2002 a 2009	NDIIPP	EUA – LC (Library of congress)	<a href="http://www.digitalpreservation.gov/multimedia/documents/ndiipp_plan.pdf">http://www.digitalpreservation.gov/multimedia/documents/ndiipp_plan.pdf</a>	Programa que involucró a cientos de organizaciones de Estados Unidos y el mundo para preservar colecciones digitales en riesgo y construir una infraestructura de preservación digital distribuida. El trabajo de NDIIPP continuó a través de la National Digital Stewardship Alliance (NDSA) mientras era administrado	Todo tipo de Documentos	NO

MODELO DE INTEGRACIÓN DE METADATOS PARA PRESERVACIÓN EN LA EDICIÓN DEL LIBRO DIGITAL

				por la Biblioteca del Congreso. A partir de enero de 2016, la Federación de Bibliotecas Digitales se convirtió en el hogar institucional de NDSA.		
2002	DPC (Digital Preservation Coalition)	Reino Unido e Irlanda	<a href="https://www.dpconline.org/">https://www.dpconline.org/</a>	Promueven, la participación de la comunidad, el desarrollo de la fuerza de trabajo, la creación de capacidad, las buenas prácticas y la buena gobernanza.	Todo tipo de Documentos	SI
2004	MetaArchive Cooperative	EUA	<a href="http://metaarchive.org/">http://metaarchive.org/</a>	Fomenta la comprensión de los métodos de preservación digital distribuida con "archivos oscuros" duraderos y estables, geográficamente dispersos.	Documentos de Archivo	SI
2006 a 2009	DPE (Digital Preservation Europe)	Comunidad Europea	<a href="http://www.digitalpreservationeurope.eu/">http://www.digitalpreservationeurope.eu/</a>	Fue un consorcio de instituciones académicas y culturales europeas dedicadas a compartir sus conocimientos y recursos colectivos de preservación digital. Financiado por la Comisión Europea.	Todo tipo de Documentos	NO
2009	DCE (The Digital Curation Exchange)	EUA	<a href="http://digitalcurationexchange.org/">http://digitalcurationexchange.org/</a>	Sitio de interacción para la comunidad de curación digital. Espacio para las discusiones sobre temas relacionados con la curación digital.	Todo tipo de Documentos	SI
2010	NDSA	EUA	<a href="http://nds.org/">http://nds.org/</a>	Consortio de organizaciones de EUA comprometidas con la preservación a largo plazo de la información digital en beneficio de las generaciones presentes y futuras. Desde 2016 pertenece a la Federación de Bibliotecas Digitales (DLF), en el Consejo de Biblioteca y Recursos de Información (CLIR).	Todo tipo de Documentos	SI
2010	OPF (Open Preservation Foundation)	Comunidad Europea	<a href="http://openpreservation.org/">http://openpreservation.org/</a>	La Open Preservation Foundation brinda a sus miembros soluciones confiables para los desafíos de la preservación digital. Administra la cartera de software de preservación digital de código abierto y permite el desarrollo de mejores prácticas a través de grupos de interés, eventos comunitarios y capacitación.	Todo tipo de Documentos	SI
2011	BitCurator Consortium	EUA (De carácter Internacional)	<a href="https://bitcuratorconsortium.org/">https://bitcuratorconsortium.org/</a> <a href="http://bitcurator.net/">http://bitcurator.net/</a>	Grupo que atiende BitCurator, es un entorno con herramientas de código abierto adaptadas de la industria forense digital para bibliotecas, archivos y museos. Las herramientas extraen objetos digitales, crean metadatos, aseguran la integridad e identifican datos confidenciales	Todo tipo de Documentos	SI
2011	DCC (The Digital Curation Centre)	Reino Unido	<a href="http://www.dcc.ac.uk/">http://www.dcc.ac.uk/</a>	Centro de conocimiento reconocido en curación digital, enfocado a la construcción de capacidades y habilidades para la gestión de datos de investigación. Brinda asesoría a organizaciones de investigación que desean almacenar, administrar, proteger y compartir datos de investigación digital.	Documentos de investigación	SI
2012	Digital POWRR (Preserving digital Objects with Restricted Resources)	EUA	<a href="http://digitalpowrr.niu.edu/">http://digitalpowrr.niu.edu/</a>	Apoya a instituciones de EU que tienen pocos recursos para hacer preservación digital.	Todo tipo de Documentos	SI

2014	LPC (Library Publishing Coalition)	EUA	<a href="https://librarypublishing.org/">https://librarypublishing.org/</a>	Es una comunidad dedicada al campo de la publicación de bibliotecas que carecen de un espacio central donde puedan reunirse, trabajar juntos, compartir información y enfrentar problemas comunes.	Editores de libros	SI
2017	PUBLISHING @W3C	EUA	<a href="https://www.w3.org/publishing/">https://www.w3.org/publishing/</a>	El grupo de trabajo EPUB 3 está estrechamente relacionado con el grupo de la comunidad editorial. Mientras que el grupo de trabajo EPUB 3 es el guardián del estándar EPUB, el grupo comunitario es el hogar de discusiones sobre nuevas direcciones técnicas, ideas y experimentación; Es el principal "brazo de incubación" o el de Publishing @ W3C, el lugar de nacimiento de nuevas ideas para el futuro.	Publicaciones electrónicas	SI
Fuente: La autora, febrero 2018.						

Como ya se dijo, todas estas iniciativas y organizaciones han generado diversos productos que apoyan el desarrollo de la preservación digital, entre los cuales se han mencionado los niveles LoP de NDSA y el sitio COPTR, ambos instrumentos que orientan a quienes se interesan en iniciar actividades de preservación. Además, existen herramientas como JHOVE que permite identificar, validar y caracterizar formatos de objetos digitales, a partir de una red de datos distribuida GDFR<sup>99</sup> e impulsada por NDSA, en la que se registran de forma cooperativa cada formato de archivo que va surgiendo con los cambios en la tecnología, entre muchos otros más.

Todos estos instrumentos y herramientas son de gran utilidad en la preservación de libros digitales, sin embargo, este trabajo de investigación no pretende describirlos ya que no es su objetivo, a pesar de esto, para quienes deseen ahondar en ese tema, se sugiere revisar COPTR, pues ahí se encuentran registradas, hasta el momento, todas las herramientas útiles en la preservación digital.

Sin embargo, además de las herramientas para la preservación digital, se han desarrollado modelos que orientan en la aplicación y desarrollo de la actividad de preservación en el ámbito digital, los cuales, si son tema de esta investigación, de ahí que a continuación se describe el Modelo OAIS y otras perspectivas.

### 1.6.2. MODELO OAIS Y OTRAS PERSPECTIVAS

El modelo OAIS fue creado por el CCSDS<sup>100</sup> de la NASA y establecido como el estándar OAIS<sup>101</sup> ISO 14721:2012 que se utiliza como orientador en la construcción de un Sistema de

<sup>99</sup> Global Digital Format Registry creada por Harvard University Library

<sup>100</sup> Consultative Committee for Space Data Systems (CCSDS) (Comité Consultivo de Sistemas de Datos Espaciales).

<sup>101</sup> Open Archival Information System. La primera versión en 2002 y la última en 2012. Es una norma de naturaleza descriptiva. El CCSDS, es un foro de agencias espaciales interesadas en el desarrollo cooperativo de normas para el tratamiento de datos para el apoyo de la investigación espacial y desarrolló estándares formales para el almacenamiento a largo plazo de datos digitales generados a partir de misiones espaciales. Sin embargo, el modelo se extiende mucho más allá de la comunidad de datos espaciales.

Archivo de Información<sup>102</sup>, principalmente para la preservación de elementos de archivo administrativo, sin embargo, su especificación sirve como referente para cualquier organización responsable del cuidado o protección de otros tipos de objetos digitales incluidos los libros.

En este estándar se define al “archivo” como una institución dispuesta a preservar, de ahí que la palabra se usa de forma general y va más allá del ámbito archivístico:

[...] un Archivo, [...], puede ser parte de una organización más grande de personas y sistemas, acepta la responsabilidad de preservar información y ponerla a disposición<sup>103</sup> [...] (CCSDS, 2012, pp. 1-1)

La palabra “archivo” en el modelo OAIS describe a la dependencia responsable del resguardo de documentos, la cual puede ser una biblioteca, que tiene establecida como función el almacenamiento, operación y preservación de objetos de información y para la cual, el modelo considera que es necesario contemplar una estructura sistémica con las siguientes acciones (OCLC & Lavoie, 2014, p. 7) (CCSDS, 2012, pp. 3-1):

- La negociación con quienes producen los objetos de información, para el ingreso apropiado al archivo.
- Alcanzar el nivel suficiente de descripción de cada objeto de información archivado, para garantizar el control y preservación a largo plazo.
- Determinar la organización de los objetos de información, de acuerdo con la comunidad a la que va dirigida.
- Asegurar la comprensión de los objetos de información.
- Seguir políticas y procedimientos documentados, para garantizar la preservación contra todas las contingencias razonables, incluida la desaparición del Archivo. No debe permitirse el borrado de objetos de información, sin una estrategia controlada.
- Permitir la consulta de los objetos de información, con la posibilidad de rastrear su autenticidad.

Así entonces, el modelo define un Sistema de Archivo de Información, el cual se refiere a una organización encargada de preservar información para lo cual, dentro de esta norma, se contempla seis funciones principales (CCSDS, 2012, pp. 4-1):

2. **Ingesta:** Se refiere a los servicios y acciones para aceptar paquetes de información y prepararlos para almacenamiento y administración dentro del sistema de archivo. Las funciones contemplan recibir y asegurarla calidad de los objetos y sus descripciones, para armar paquetes y enviar al repositorio de archivado permanente que forma parte del mismo sistema de archivo.

---

<sup>102</sup>Open Archival Information System (OAIS). Usa el término "Abierto" para hacer énfasis en el hecho de que el modelo fue desarrollado y lanzado en un foro público abierto, en el cual participó todo el que estuviera interesado, por lo tanto, NO implica un acceso irrestricto a los objetos de información.

<sup>103</sup>Traducido por Ana Yuri Ramírez

3. **Repositorio:** Proporciona los servicios y funciones para el almacenamiento, mantenimiento y recuperación de los paquetes archivados. Las funciones incluyen: recibir en el almacenamiento permanente, administrar las estructuras de almacenamiento, realizar control de existencias (comprobaciones rutinarias y especiales de errores y establecer capacidades de recuperación de desastres) y atender las solicitudes de acceso a los paquetes.
4. **Gestión de datos:** Se refiere a los servicios y funciones para generar, mantener y acceder a la información descriptiva y administrativa que identifica y documenta a los paquetes y que es utilizada para administrar el Archivo. Las funciones incluyen: administración de la base de datos (mantener su estructura e integridad), actualización de la base de datos (carga y actualización de información) y consultas sobre la base de datos.
5. **Administración:** Se refiere a los servicios y acciones relacionadas con el funcionamiento general del Sistema de Archivo. Las funciones incluyen: solicitar y negociar las entregas y auditar las para que cumplan con los estándares del Archivo, dar mantenimiento al hardware y software, supervisar el sistema, hacer mejorar continua en las operaciones del Archivo (inventario, informes, migraciones y actualizaciones de los contenidos del archivo), establecer y mantener estándares y políticas para el Archivo y proporcionar soporte al usuario de consulta.
6. **Plan de preservación:** se refiere a los servicios y funciones para monitorear el entorno del OAIS, hacer recomendaciones y planes de preservación para garantizar el acceso y comprensión del acervo, incluso si la infraestructura tecnológica original se vuelve obsoleta. Esto incluye: evaluar los contenidos (curaduría) <sup>104</sup> para actualizar periódicamente y migrar, desarrollar recomendaciones de estándares y políticas, análisis periódico de riesgos e informes (monitorear los cambios en la tecnología y los requisitos de servicio de los usuarios). Además, diseñar plantillas de paquetes de información y planes de migración detallados (software, prototipos y pruebas) para aplicarlos en la administración del OAIS.
7. **Acceso:** Proporciona los servicios y funciones de consulta, determina la existencia, descripción, ubicación y disponibilidad de la información almacenada en el OAIS. Incluye: recibir solicitudes, controles para limitar el acceso a la información protegida, control para la ejecución exitosa de las consultas.

Estas funciones interactúan de acuerdo con un esquema, en donde la comunicación es a través de paquetes de información<sup>105</sup>, los cuales adquieren características especiales de acuerdo con el flujo en el que se encuentran y que a continuación se describen, ver Figura 5, (CCSDS, 2012):

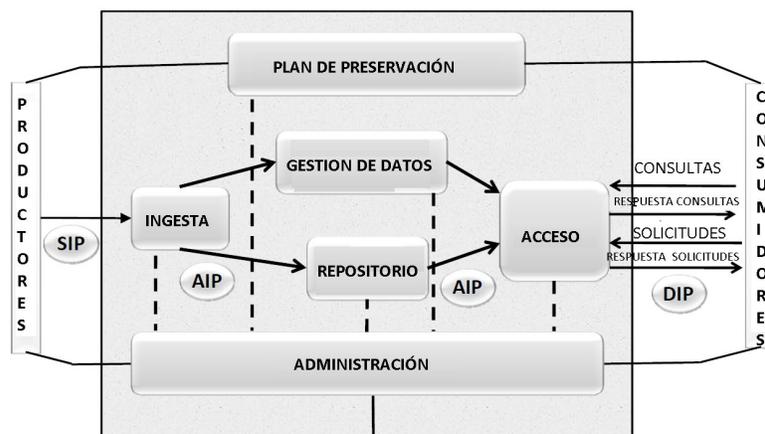
---

<sup>104</sup> Implica mantener, preservar y agregar valor a los objetos digitales a lo largo de su ciclo de vida. (DCC, 2004)

<sup>105</sup> Contenedor lógico con información descriptiva del objeto y su contenido e información asociada con el proceso de preservación

- **SIP** (Submission Information Package): Paquete de información solicitado, es lo entregado por el creador del objeto digital, que debe incluir información descriptiva además del contenido.
- **AIP** (Archival Information Package): Paquete de información para archivo, consiste en el contenido del SIP y la información generada para preservación, que permite conservarlo dentro del OAIS.
- **DIP** (Dissemination Information Package): Paquete de información de distribución, es derivado de uno o más AIP y enviado al consumidor en respuesta a una solicitud específica al OAIS.

Figura 5. Modelo OAIS



Fuente: CCSDS, 2012

Este modelo está basado en un conjunto de funciones que parten de la recepción de objetos digitales, empaquetados en una estructura determinada por el OAIS y derivados de una entidad (productores), los cuales son tratados de forma especial para poder conservarlos dentro del Sistema de Archivo y así, proporcionar servicios que atiendan las solicitudes o consultas por parte de entidades interesada en la información (consumidores).

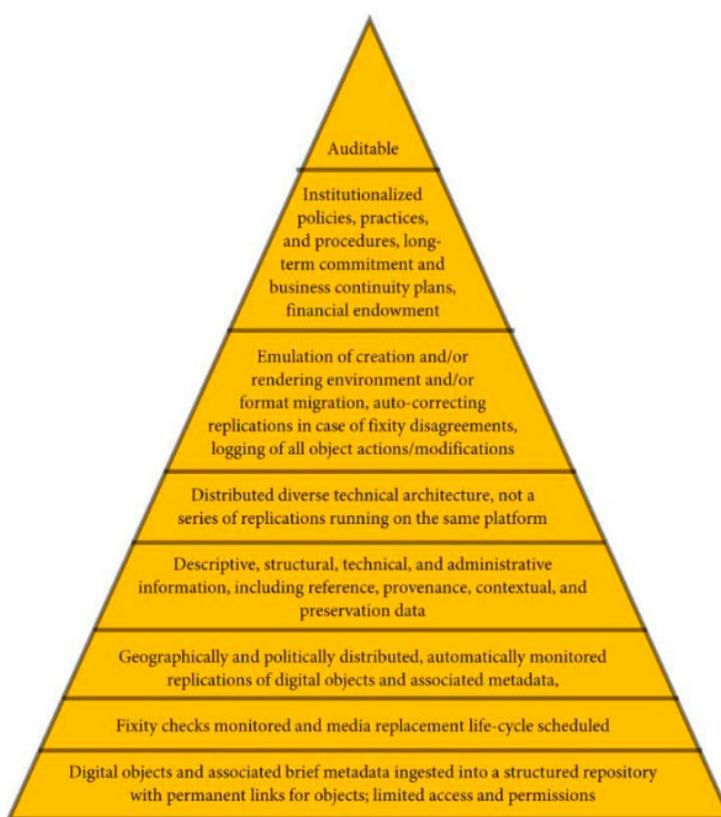
Sin embargo, las características precisas de estos elementos no están especificadas en el modelo, por lo cual, es necesario apoyarse de otros instrumentos para comprender por dónde comenzar y validar que se está haciendo lo adecuado, ya que, la cantidad de funciones que deben cumplirse demandan una inversión de recursos son muy alta, y de no hacerlo, podría ser un desperdicio de esfuerzos.

Por todo lo anterior, una la institución interesada en la preservación digital debe contar con herramientas tecnológicas y una estructura organizacional apta para el desarrollo de normas y políticas que apoyen la ejecución y priorización de las funciones establecidas en OAIS, de ahí que, Wilson señala que para llevar a cabo la preservación digital no basta con adquirir una herramienta tecnológica (Wilson, 2017) o conocer el modelo OAIS, además es

necesario entender cuáles son las necesidades básicas de la preservación y así, establecer una nueva organización.

En ese sentido, Thomas Wilson, especialista en tecnologías, bibliotecas y profesor de la universidad de Alabama, propone una *jerarquía de necesidades básicas para la preservación digital*, semejante a una pirámide de Maslow<sup>106</sup>, en la cual se describen y caracterizan algunos puntos que ya son mencionados dentro de las funciones del modelo OAIS, sin embargo, en la propuesta se les da un peso o valor, pues de acuerdo con el autor existen acciones que mitigan un mayor porcentaje de riesgos asociados con los objetos digitales, por lo cual, tomar como guía esta pirámide para priorizar las acciones al definir una estrategia de preservación, ayudaría a darle solidez, ver Figura 6.

Figura 6. Jerarquía de necesidades de preservación digital.



Fuente: Wilson, 2017.

De acuerdo con el autor, cada nivel en la pirámide determina la prioridad o importancia que el elemento descrito adquiere, de ahí la similitud con la pirámide de Maslow, por lo que la satisfacción es de forma jerárquica, de abajo hacia arriba, dirigiendo esfuerzos de tal forma

<sup>106</sup> Abraham Maslow: Teoría de la motivación y el desarrollo humano, expresada en una pirámide de jerarquía de necesidades que comienza con las necesidades fisiológicas básicas y termina con autorrealización. Cada peldaño de la pirámide requiere el cumplimiento del anterior (Wilson, 2017).

que lo importante es satisfacer las necesidades básicas o primordiales y así, ir reduciendo los riesgos básicos de pérdida en los objetos, hasta llegar a la última posición que es la ideal.

De lo anterior que, para Wilson, el contar con una amplia cantidad de metadatos está en la posición más baja de su pirámide, de ahí que es el elemento primordial para lograr la preservación digital y sin esto, desde su perspectiva no es posible alcanzar los siguientes niveles.

En concordancia con Wilson, Angela Dappert y Markus Enders, especialista en metadatos de la biblioteca británica, se refieren a los metadatos o información referente a los objetos almacenados, como un requisito determinante para la conservación de los mismos, pues desde su perspectiva estos permiten (Dappert & Enders, 2010):

- Tener objetos digitales ejecutables, enteros, intactos y originales de forma permanente.
- Contar con la identificación única y persistente del objeto digital.
- Claridad en cómo se relacionan todas las partes que componen al objeto digital.
- Conservar las características significativas de los objetos, incluso si los soportes o los medios físicos cambian.
- Contar con la información que permita entender al objeto digital disponible.
- Velar por el cuidado de los medios físicos.
- Tener el control físico del repositorio o contenedor de objetos.

De lo anterior, que lo señalado por Dappert y Enders concuerda con lo señalado en la pirámide de Wilson, ya que ellos también consideran como parte fundamental de la preservación digital, la adecuada ingesta de los objetos digitales en un repositorio, contemplando metadatos asociados.

En conclusión, es indispensable remarcar que tanto el modelo OAIS con la perspectiva de Wilson, y lo señalado por Dappert y Enders concuerdan con la importancia de los metadatos, de ahí que, a pesar de que dichas propuestas no fueron elaboradas específicamente para aplicarse en libros digitales, para fines de esta investigación se tomarán como punto de referencia, ya que son aplicables en lo general para la preservación digital y son el punto de partida para tratar el siguiente tema, los metadatos.

---

### 1.6.3. LA PRESERVACIÓN DEL LIBRO DIGITAL

Antes de entrar de lleno, en la importancia de los metadatos para la preservación digital del libro, conviene puntualizar algunos otros aspectos importantes, señalados por algunas de las organizaciones antes listadas, en algunos casos, como aspectos que afectan a todo objeto digital, incluido el libro.

De lo anterior que, los peligros más importante que toda institución debe considerar y combatir, en la búsqueda de proteger a cualquier recurso digital fueron especificaros por primera vez, por el NDIIPP<sup>107</sup> (Harvey, 2015) y se refieren a:

- **Peligros tecnológicos:** Agrupa los retos relacionados con mantener el hardware y software utilizable para la lectura de un objeto digital e incluye los problemas relacionados con la amplia diversidad de formatos de archivo, que vuelven frágil al contenido, ya que éste puede quedar inutilizable y/o degradarse en su composición de bits, por los cambios en los medios de almacenamiento y programas de cómputo.
- **Peligros legales y políticos:** En este ámbito se encuentra la falta de claridad o respaldo en las leyes, por un lado, los derechos de autor, que en algunos casos prohíben o se contraponen a la recopilación de la producción digital, y por otro, las leyes de Depósito Legal que carecen del reconocimiento público del valor del contenido digital, dificultando la recopilación de estos materiales.
- **Peligros de contenido:** Estos se refiere a tres factores importantes, primero la rapidez con la que crece la producción de materiales digitales, segundo el incremento en la complejidad de éstos derivada de los avances tecnológicos, y por último, la falta de documentación en el objeto, que es indispensable para almacenar, organizar y recuperar y que, por su ausencia, es costosa y muchas veces difícil de generar.
- **Peligros organizacionales:** Se tiene identificado que estos procesos requieren de una fuerte inversión de recursos tecnológicos, que difícilmente una sola institución puede soportar, por lo que se advierte la necesidad de trabajar de forma colaborativa e indispensable compartir y transmitir dichas experiencias, ya que siempre implican recursos y personal calificado.

Adicionalmente, respecto a la preservación del libro digital, DPC señala que existen los siguientes riesgos por combatir, con base en el tipo de contenido o estructura con la que se construye el libro (Kirchhoff & Morrissey, 2014):

- **Por digitalización:** Este tipo de libros es muy popular debido a los múltiples programas de digitalización en el mundo, sin embargo, su acelerado crecimiento genera problemas en su producción ya que en su mayoría carecen de metadatos y OCR, además de omitir el uso de formatos de imagen y archivo estandarizados, está la falta de un esquema de almacenamiento ordenado y estructurado.
- **Como espacios de referencia:** Son materiales que pueden ser publicados desde otra perspectiva comercial al libro, como enciclopedias, diccionarios, bibliografías y que generalmente están en bases de datos y en constante actualización, por lo cual las instituciones requieren establecer formas adecuadas de preservación similares a las acciones utilizadas tradicionalmente para el mantenimiento de base de datos o web dinámicos.

---

<sup>107</sup> Creado en 1998 por la Biblioteca del Congreso de los Estados Unidos de América (LC),

- **Como contenido con enlaces embebidos:** Este tipo de libros se caracteriza porque en su estructura contemplan objetos digitales adicionales como audios, videos y enlaces a otros contenidos, que por un lado tienen sus propios derechos asociados y además pueden estar dispersos en otros lugares virtuales. Por lo que es necesario establecer los mecanismos para el alcance de dichos objetos, tanto en la especificación de derechos de autor para incluir a los objetos asociados y en las estrategias de tecnología para recolectar los elementos que pueden estar ligados y almacenados en espacios externos al libro.
- **Como textos referenciados:** Se refiere a los contenidos textuales que generalmente se encuentran embebidos o forman parte de otros tipos de objetos, como cursos en línea u otra publicación digital y que deben ser identificados para ser preservados de forma individual como libros.

Además, la DPC define en siete rubros a los principales problemas, que las instituciones deben enfrentar específicamente, para llevar a cabo la preservación del libro digital (Kirchhoff & Morrissey, 2014):

- **El Depósito Legal:** Existen muchas naciones que exigen a sus editores el depósito de sus publicaciones en una Biblioteca Nacional, sin embargo, las modificaciones en las Leyes de Depósito Legal para integrar a los libros digitales son recientes y en cada nación hay condiciones propias. De ahí que, no todas las Bibliotecas Nacionales tienen el instrumento legal adecuado para preservar los libros digitales para sus ciudadanos y hacerlos accesibles.
- **La auto-publicación:** Existen novedosas herramientas<sup>108</sup> que permiten la edición y publicación de forma personal de un libro, por lo cual esta práctica se ha incrementado y hace más compleja la labor de las bibliotecas para la identificación y ubicación de nuevas obras, ya que este tipo de obras están fuera de los registros de derechos de autor e ISBN, por lo tanto, su recuperación y preservación en muchos casos es imposible.
- **Estabilidad del contenido:** Las obras de consulta generalmente se construyen con la facilidad de editar y publicar de ahí que su actualización se hace con facilidad e intencionalmente de forma discreta y silenciosa, por lo que es difícil saber cuándo una nueva versión de un libro o publicación requiere ser preservada.
- **El modelo de negocio del libro:** Ha cambiado, pues los tirajes de impresión y la distribución son bajo demanda, la venta se ha transformado en suscripciones de acceso y la intermediación entre el editor y el lector se ha reducido, además los editores requieren nuevas habilidades y conocimientos en tecnología.
- **Permisos o derechos, DRM:** Se refiere a la tecnología empleada para proteger los derechos de autor de los contenidos digitales que puede funcionar de varias formas, por ejemplo, restringiendo la frecuencia o duración del acceso al material o las

<sup>108</sup> Kindle Direct Publishing, NOOKPress, iBooks Autor

posibilidades de copiado y transferencia, además puede ayudar a evitar alteraciones en el objeto. Sin embargo, la problemática principal para las instituciones encargada de la Preservación Digital es que este tipo de tecnología hace inteligible el objeto, de tal forma que es posible preservar la cadena de bits, pero el contenido para la consulta es inaccesible.

- **El no acceso al libro para las bibliotecas:** Como ya se dijo, el libro ahora se comercializa como servicios de suscripción y se protegen los derechos de los autores con los DRM<sup>109</sup> sin embargo, esto a pesar de beneficiar al libro, en la mayoría de las veces impide a las bibliotecas poseer el libro digital, lo cual afecta directamente en la preservación, pues las bibliotecas no tienen el libro, ni los permisos para garantizar su permanencia.
- **La propiedad del libro:** Como ya se mencionó, la nueva forma de comercializar los libros electrónicos está cambiando el sentido de propiedad del libro, pues antes con el libro impreso el lector adquiría la libertad de hacer con él lo que quisiera, sin embargo con el libro digital no es así y en muchos casos las empresas sólo venden licencias de uso, cuyas políticas pueden ir desde aquellas en la que el libro se queda almacenado en un espacio virtual, sin que sea descargable y el lector sólo tiene la posibilidad de leerlo siempre que mantenga perpetuamente el pago del servicio, hasta aquellas en las que el libro se descarga en un dispositivo específico pero no es posible imprimirlo o compartirlo con alguien más.

En conclusión, de este apartado, la preservación digital para el libro digital debe atenderse desde una perspectiva general en cuanto a los riesgos que corre por ser un objeto digital, pero además tiene aspectos particulares que tienen que ver con su construcción y los cambios que deben atenderse en cuanto a las legislaciones y las nuevas formas en su construcción o modelo de negocio.

Además, la forma en que se organiza la biblioteca (modelo OAIS) para establecer una relación con los editores será importante, y debe comenzar por establecer una conformación basta de metadatos, de ahí que serán importante conocer cuál es el modelo que los editores emplean para la edición del libro y cuál es la cantidad de metadatos que se deben generar respecto al libro, de ahí que son los temas que a continuación se abordarán.

### 1.7. METADATOS PARA PRESERVACIÓN Y DE PRESERVACIÓN.

En este apartado se hablará primero del concepto de metadatos para después entender cómo se clasifican y a partir de esto establecer lo que, para fines de esta investigación, se llamarán **metadatos para preservación** y la diferencia de estos, con respecto a los que se definen como **metadatos de preservación**.

---

<sup>109</sup> Digital Rights Management.

### 1.7.1. EL CONCEPTO DE METADATO

La palabra metadato o en inglés metadata, se conceptualiza como “datos acerca de datos” (Howe, 2010) pues se refiere a “información descriptiva sobre el contexto, calidad y condición o características del dato” (Senso & Rosa Piñero, 2003), esto se respalda con la definición de la palabra dato, desde el ámbito de la computación:

[...] podemos decir que los datos vienen en tres tipos: como datos que se refieren a algo fuera de sí mismo ... los datos como datos, ... y datos que funcionan en datos [...] <sup>110</sup> (Ryan et al., 2014, p. 125)

Sin embargo, Pomerantz considera que esta primera definición carece de claridad, pues desde su punto de vista, el dato es información en bruto y sin procesar y el metadato no, ya que, de entrada, proporciona información, por lo que prefiere referirse a él como:

[...] un objeto potencialmente informativo que describe a otro objeto potencialmente informativo [...] <sup>111</sup> (Pomerantz, 2015, p. 26)

Esta definición concuerda con la perspectiva establecida desde el ámbito de la computación, que define al metadato como:

Datos estructurados sobre un objeto que admite funciones asociadas con el objeto designado <sup>112</sup> (Sicilia & World Scientific (Firm), 2014)

En esta definición de Sicilia se inserta una característica importante, la estructura, que conlleva organización, uniformidad y orden en los datos, con lo cual se convierten en un objeto informativo y deja de ser un dato sin procesar como lo dice Pomerantz, de tal manera que al hablar de metadatos se incluye un proceso que implica la organización de los datos en esquemas o estructuras que den orden y sentido para generar información.

Este proceso puede verse como lo señaló Ercegovac en 1999, a partir del elemento que se describe como un objeto bibliográfico, registro archivístico, recurso visual, objeto geoespacial o incluso productos de software, en donde cada uno de estos presentar un nivel diferente de especificidad y estructura, en cuanto al proceso de generación de datos, sin embargo la intención del metadato se centra en describir, identificar y definir el recurso para posteriormente poder recuperarlo, autenticarlo, evaluarlo e informar acerca de sus condiciones de uso, interoperabilidad y preservación (Senso & Rosa Piñero, 2003).

De ahí que vale la pena mencionar cuáles son las características que deben cubrir los metadatos para preservación según la UNESCO (UNESCO & BNA, 2003, p. 100):

- Definen al material bajo la responsabilidad de un programa de preservación.
- Indican qué se necesita para mantener y proteger.

<sup>110</sup> Traducido por Ana Yuri

<sup>111</sup> Traducido por Ana Yuri

<sup>112</sup> Traducido por Ana Yuri

- Indica qué se necesita para representar al objeto cuando el usuario lo solicite o cuando las tecnologías de almacenamiento y acceso hayan cambiado.
- Registra la historia del objeto y las consecuencias de lo que le vaya sucediendo.
- Documenta la identidad e integridad del objeto para certificar su autenticidad.
- Permite al usuario y al programa de preservación comprender el contexto del objeto almacenado y en uso.

Además, como ya se dijo, la generación de metadatos implica procesos de ahí su realación con la gestión documental, ya que se habla de describir e identificar para después recuperar, lo cual claramente se especifica en la norma ISO 15489 que, además, también define dentro su ámbito al concepto de metadatos como:

datos que describen el contexto, contenido y estructura de los documentos, así como su gestión a lo largo del tiempo (ISO 15489-1:2001, 3.12 citado por REDC, 2008).

En este mismo sentido, la norma ISO 23081-1:2006 que se refiere específicamente a los metadatos para la gestión de documentos, los define como:

[...] información estructurada o semiestructurada que posibilita la creación, registro, clasificación, acceso, conservación y disposición de los documentos a lo largo del tiempo y dentro de un mismo dominio o entre dominios diferentes [...] (REDC, 2008, p. 276).

Esta última norma, en su definición además señala dos momentos importantes dentro de un sistema de gestión documental para los metadatos, el primero que abarca la cantidad de información requerida para la incorporación de un documento en un ambiente de gestión y el segundo, que se refiera a los metadatos generados una vez incorporado el documento en el sistema de gestión.

Para la incorporación en el sistema de gestión documental, señala que los metadatos abarcan elementos del contexto en el que es creado el documento, es decir, la organización e involucrados, además del contenido, apariencia, estructura y atributos técnicos, en donde el contexto también debe ser generado por el productor pues incluye información del proceso de creación y origen del documento, además de la estructura y atributos técnicos, que en esencia debe contemplar dos elementos (REDC, 2008, p. 279):

- a) su estructura física o técnica, y
- b) su estructura lógica, esto es, las relaciones entre los elementos de datos que lo componen

El segundo instante, se refiere a los metadatos que acompañaran al documento una vez contenido en un sistema de gestión documental y que se derivan de los procesos aplicados, también conocidos como *metadatos sobre actividades de gestión*, que documentan acciones aplicadas al documento para garantizar su autenticidad, integridad, fiabilidad y disponibilidad a lo largo del tiempo y que incluyen, metadatos técnicos, cambios en el contenido, contexto y estructura, acciones administrativas como acceso al documento, los

metadatos de ingreso e incluso la información registrada cuando un documento se migra a otro sistema de gestión, para justificar su inexistencia.

En este sentido, vale la pena retomar al Modelo OAIS para analizar el modelo de información que maneja como una guía para un sistema de gestión de archivo para preservación y que en estricto sentido es también un sistema de gestión documental, como el señalado en la norma ISO 15489, y que en OAIS se especifica como la Información Empaquetada (**IP**)<sup>113</sup>, descrita como un contenedor conceptual con dos elementos informativos (CCSDS, 2012):

- **Información de contenido (CI)**<sup>114</sup>. La cual, se compone por dos elementos: Un objeto dato contenido (físico o lógico) y la información de representación necesaria (metadatos) para comprender al objeto por quienes lo consultan o consumen.
- **Información de descripción de preservación (PDI)**<sup>115</sup>. La cual, está definida con base en el contenido del CI, sirve para identificarlo claramente, comprender el medio en el que se creó para preservarlo así, con base en cinco elementos informativos (metadatos): procedencia, contexto, referencia, permanencia y derechos de acceso.

Luego entonces, si se compara lo señalado en la norma ISO 23081-1:2006 con este paquete IP de OAIS, concuerdan tanto en los metadatos recibidos o generados en la ingesta y que ayudan a comprender al objeto (CI) como en los *metadatos sobre actividades de gestión* (PDI) que se refieren a los procesos aplicados dentro de un sistema OAIS, en donde este bloque, a su vez se constituye de cinco elementos descriptivos referentes a los procesos de preservación definidos en modelo OAIS de la siguiente manera (CCSDS, 2012):

- **Procedencia.** Datos acerca del origen del CI, nacimiento, custodia e historial que incluya los procesos por los que ha pasado.
- **Contexto.** Datos para establecer relaciones del CI con otros paquetes de información o elementos externos que permitan enriquecer el origen del CI.
- **Referencia.** Datos para identificar de forma única al CI.
- **Permanencia.** Datos que van a proporcionar una envoltura o escudo protector para el CI, pues permiten identificar cualquier alteración no programada. Como el registro de la composición de la cadena de bits, de cada archivo de computadora.
- **Derechos de acceso.** Datos que permiten ejecutar las maniobras adecuadas para el acceso al CI, tanto en las acciones de preservación (migraciones o transformaciones) como en las consultas, para respetar los derechos de distribución establecidos por el creador.

---

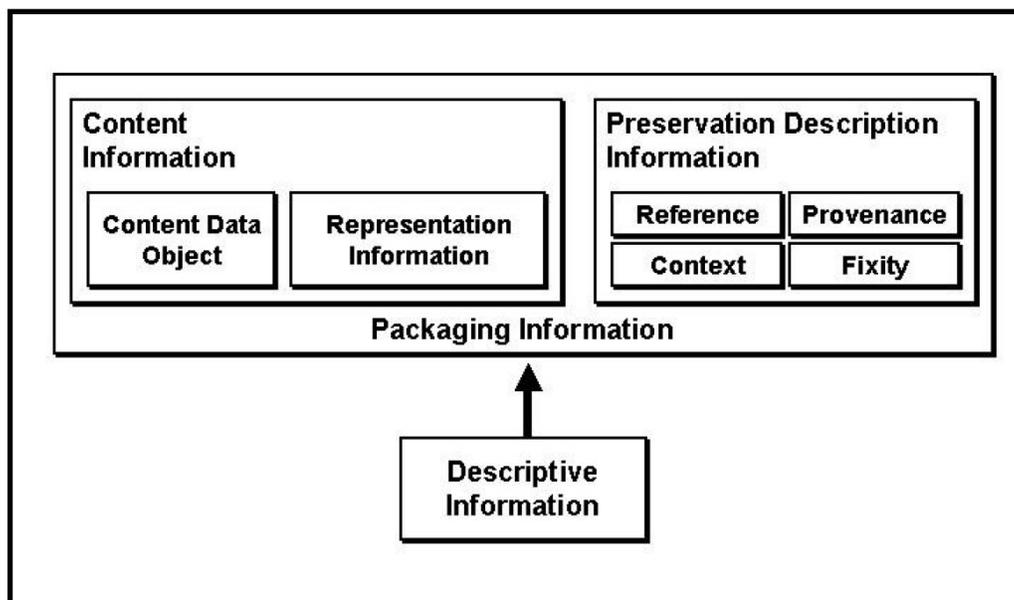
<sup>113</sup> Packaging Information

<sup>114</sup> Content Information

<sup>115</sup> Preservation Description Information

Además, en el modelo de Información OAIS, existe la Información Descriptiva (DI)<sup>116</sup> del paquete de información, la cual sirve dentro del sistema de gestión OAIS para identificar qué está contenido en el IP. Es importante recalcar que esta información (metadatos) puede ser una réplica del contenido completo del paquete, pues está determinado de acuerdo con las necesidades de la entidad que implementa el OAIS, ver Figura 7.

Figura 7. Modelo de Información OAIS.



Fuente: DPC et al., 2005.

Como puede verse el modelo de información OAIS concuerda con la visión de los metadatos en la norma ISO 23081-1:2006, pues contempla tanto los metadatos que se generan desde la ingesta y aquellos resultados de la gestión, sin embargo, ninguna de las dos normas especifica la estructura de estos, pues existen diversas unidades de información, por lo que, se han desarrollado varias propuestas para la definición de estas (DPC et al., 2005), de ahí que a continuación se revisarán algunas propuestas de diversos autores y organizaciones.

Guilliland en 1998 propuso las siguientes categorías para la organización de clasificación de los metadatos (Senso & Rosa Piñero, 2003):

- **Administrativos:** Datos usados en la gestión y administración de objetos digitales.
- **Descriptivos:** Datos representativos del objeto digital
- **Preservación:** Datos que permiten salvaguardar al recurso digital.
- **Técnicos:** Datos relativos al funcionamiento o comportamiento del objeto digital.
- **Uso:** Datos relativos a los permisos de uso del recurso digital.

<sup>116</sup> Descriptive Information.

Por su parte, el Consejo Internacional de Archivos (ICA)<sup>117</sup> propone otra tipología (Candás-Romero, 2006):

- **Descriptivos:** Son los datos generados por la misma organización creadora del objeto digital (autor, fecha, título, etc.).
- **Archivo:** Son los datos añadidos para facilitar la gestión del objeto (la última revisión, fecha de recepción en el archivo, condiciones de uso, etc.)
- **Técnicos:** Datos necesarios para comprender y procesar al objeto (formato de archivo de computadora, fecha de la última migración de formato, etc.)
- **Premis**<sup>118</sup>: Información para apoyar el proceso de preservación digital, metadatos para el control de la viabilidad, legibilidad y accesibilidad, autenticidad e identificación en el contexto de preservación.

Dappert y Enders, mencionan que hay cuatro categorías (Dappert & Enders, 2010):

- **Descriptivos:** Representa al objeto a través de propiedades como autor y título, contexto histórico, procedencia, entre otros.
- **Estructurales:** Registra las relaciones estructurales físicas y lógicas (asociaciones con archivos y/o ligas a otros recursos).
- **Técnicos:** Información técnica que aplica a cualquier tipo de archivo de computadora, como: el software y el hardware en el que se puede ejecutar, firmas y certificados digitales de integridad y autenticidad e información respecto al tipo de contenido (ancho de la imagen, duración en el audio, tipo de codificador, etc.).
- **Administrativos:** Contempla información de procedencia, como quién se ha ocupado del objeto digital y qué acciones de preservación se han realizado en él y de derechos y permisos para el acceso al objeto, que incluye las acciones de preservación permitidas.

Cabe señalar que existen diversas propuestas para la organización de los metadatos, en las cuales algunos de los rubros antes señalados se desglosan más o subdividen principalmente en metadatos descriptivos para características de propiedad y legales, sin embargo, al comparar estas tres propuestas es posible identificar ciertas concordancias en cuanto a las clasificaciones.

Lo primero es que todos coinciden en que debe existir una información *descriptiva*, la cual hace referencia a la obra intelectual. Además, están los datos *técnicos y/o estructurales* del objeto, en donde se describe al elemento archivo de computadora, de forma individual y cuando forma parte de un conjunto de archivos con la estructura que los relaciona o

---

<sup>117</sup> International Council on Archives.

<sup>118</sup> Preservation metadata: implementation strategies.

contexto que requiere. Y en este sentido, Guillard agrega el comportamiento y funcionamiento del objeto.

También están los datos de uso o permisos de acceso que inicialmente se refieren a los derechos otorgados por el creador y en donde hay quienes también incluyen datos que tienen que ver con la autenticidad o cuestiones jurídicas, sin embargo, dentro de un sistema de gestión documental, también se relacionan con los permisos o niveles de acceso al recurso, por lo que hay quienes además los agrupan dentro de los datos administrativos, los cuales además contemplan fechas de ingreso o recepción, ubicación del recurso, cambios en el recurso y accesos.

Estos datos administrativos en algunos casos hay quienes también los relacionan con los datos de preservación como Dappert y Enders, sin embargo, hay quienes los separan ya que concretamente se refieren a las acciones encaminadas al mantenimiento a largo plazo, como migraciones y revisiones de integridad como lo hace ICA.

Así pues, para la preservación a largo plazo de un objeto es necesario contar con una variedad de metadatos que van desde los datos descriptivos, estructurales, técnicos hasta los administrativos que se generan dentro de las organizaciones a partir de los procesos ejecutados para el resguardo, mantenimiento y protección de estos, que además deben contemplar a los metadatos de preservación que específicamente se concentran en el control de la autenticidad y legibilidad del objeto.

De lo anterior, para este trabajo de investigación es muy importante distinguir entre los **metadatos para preservación** y los **metadatos de preservación**, pues el mismo modelo OAIS habla de varias unidades de información que son importantes para la preservación de un objeto digital, que como ya vimos contempla una diversidad amplia de datos, sin embargo, señala que existe un conjunto de datos que son llamados específicamente **de preservación**, por tratarse de “los metadatos que respaldan el proceso de preservación digital” (DPC et al., 2005) y permiten controlar la autenticidad, legibilidad y acceso, en donde estas dos última características se refiere específicamente a la posibilidad de visualizar el objeto en su forma original.

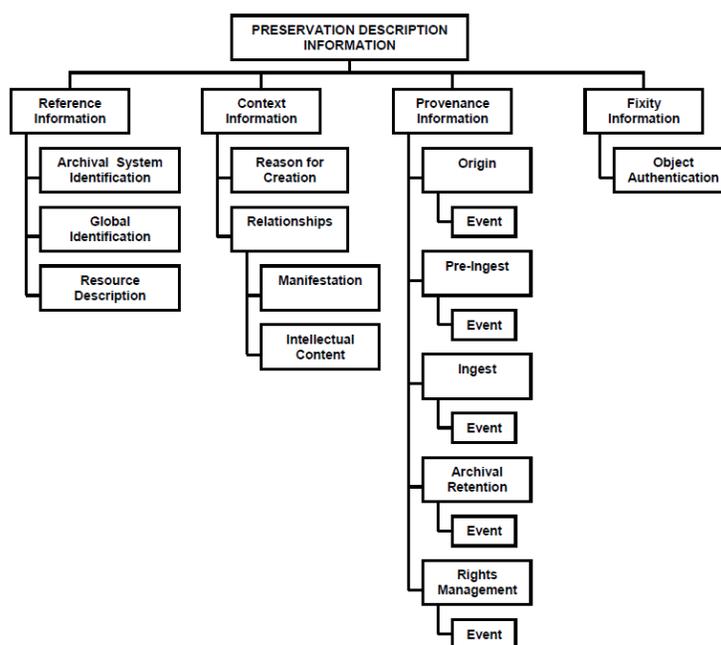
Para esto es necesario retomar lo que señala la norma ISO 23081-1:2006 cuando se refiere a la información requerida para incorporar un objeto a un sistema de gestión documental con la posibilidad de mantenerlo a largo plazo, en donde se señala información relacionada con el contexto de creación, que incluye proceso de creación y origen del objeto además del contenido, apariencia, la estructura y atributos técnicos (REDC, 2008) y que para OAIS, se concentra dentro de la Información de Contenido o CI, que además en ambos casos se recalca, es información que debe ser generada por el productor o creador de objeto. Esto es lo que, para fines de esta investigación, llamaremos **metadatos para preservación** y se refieren a **toda la información que es importante recibir o generar para que un objeto digital ingrese a un sistema de archivo** con fines de preservación.

Por otro lado, tenemos los **metadatos de preservación** que la norma ISO 23081-1:2006 los señala como toda la **información generada dentro del sistema de gestión documental para la preservación a largo plazo** y que en OAIS lo define como un modelo de información, ver Figura 5, compuesto por varios paquetes de información en la cual desde luego está la información de contenido, pero además existe un paquete de información que concentra todas las acciones de preservación o PDI.

### 1.7.2. METADATOS DE PRESERVACIÓN

Cabe señalar que en esta investigación se abordaran ambos, sin embargo, comenzaremos por describir los metadatos de preservación, los cuales de acuerdo con el modelo OAIS se dividen en: Información de referencia, Información de contenido, información de procedencia e información de integridad. Ver Figura 8.

Figura 8. Estructura completa del PDI en OAIS.



Fuente: CCSDS, 2012.

Como ya se dijo en su momento, los metadatos requieren contar con una organización y estructura, de ahí que cada categoría de metadatos requiere una especificación conocida como esquema, que ayuda a dar una forma estandarizada de interpretación a cada tipología, por lo que el esquema para los metadatos de preservación es PREMIS<sup>119</sup>, definido

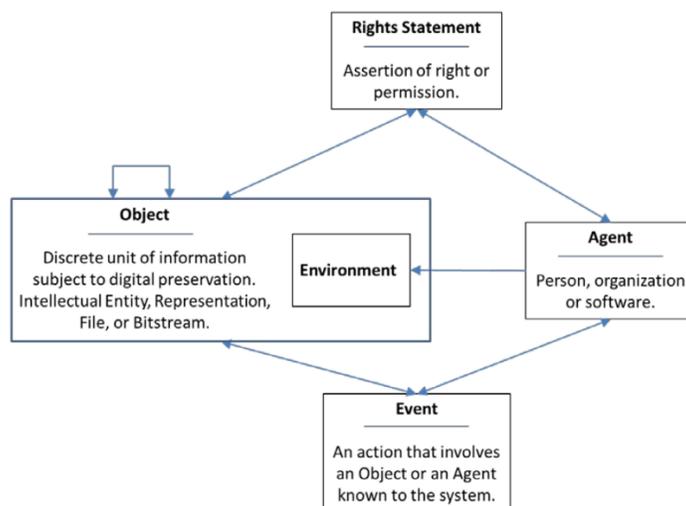
<sup>119</sup> Preservation Metadata: Implementation Strategies

en 2003 por OCLC<sup>120</sup> y RLG<sup>121</sup> compuesto de cuatro entidades<sup>122</sup>, las cuales se encuentran presentes durante todo el proceso de preservación, ver Figura 9 (PREMIS V3-LC, 2015):

1. **Objeto:** Una unidad de información discreta en forma digital, por ejemplo, un archivo PDF.
2. **Evento:** una acción de preservación, por ejemplo, ingesta del archivo PDF en el repositorio.
3. **Agente:** persona, organización o programa de software asociado a un evento, por ejemplo, el editor del archivo PDF que lo deposita en el repositorio.
4. **Derechos:** uno o más permisos pertenecientes a un Objeto, por ejemplo, permiso para hacer copias del archivo PDF con fines de preservación.

Estas entidades están definidas con base en un conjunto de datos, llamados *unidades semánticas* de información, las cuáles definen sólo qué información debe haber, de ahí que, esta puede contener un único elemento de metadatos, o dividirse en múltiples elementos de metadatos (PREMIS V3-LC, 2015).

Figura 9. Modelo PREMIS.



Fuente: PREMIS V3-LC, 2015.

Además, dentro de la unidad semántica conocida como *Objeto*, está lo que se refiere al sujeto de preservación, compuesta a su vez de cuatro subcategorías descriptivas (PREMIS V3-LC, 2015), ver Figura 10:

1. **Entidad Intelectual:** Conjunto coherente de contenido que se describe como una unidad intelectual o artística: por ejemplo, un libro. Esta puede tener una o más

<sup>120</sup> Online Computer Library Center

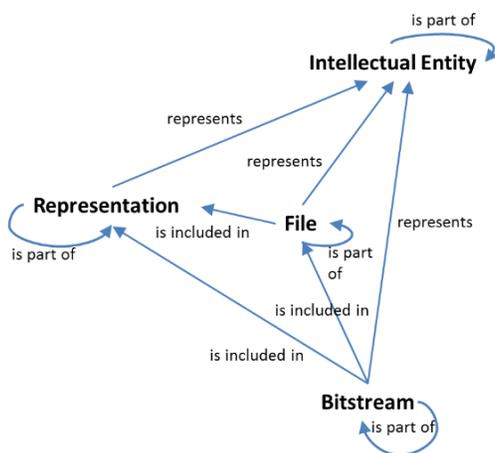
<sup>121</sup> Research Libraries Group

<sup>122</sup> Abstracción de un objeto, individuo, concepto o cosa que existe en la realidad y que por lo tanto, es independiente y puede describirse como una estructura de datos con comportamiento.

representaciones digitales o no digitales, pues se refiere a la obra como una concepción intelectual.

2. **Representación:** es el conjunto de elementos que componen la obra, para el caso de las obras digitales, se trata del conjunto de archivos y metadatos estructurales que permiten la configuración o composición de la entidad intelectual como una obra digital.
3. **Archivo:** Es una cadena de bits ordenada y agrupada bajo un mismo nombre y formato, reconocida como una unidad dentro de un sistema operativo de computadora.
4. **Cadena de bits:** Son datos contiguos o no contiguos dentro de un archivo, la cual requieren de elementos estructurales para que se asocien a un formato de archivo específico.

Figura 10. Vista conceptual entre las subcategorías que componen al objeto.



Fuente: PREMIS V3-LC, 2015.

Lo que en esencia señalan estas subdivisiones del objeto dentro de PREMIS son las distintas perspectivas desde las cuales se debe describir al objeto digital, de ahí que se puede representar como como una entidad intelectual, es decir la obra o creación conceptual, un archivo de computadora con características referentes a su formato o tipo de archivo o una cadena de bits, codificada en un hardware, pero además, el objeto digital puede tener una representación o varias, compuesta a su vez por uno o varios archivos o cadenas de bits.

Es decir que un libro digital podría ser un solo archivo PDF o un epub que se compone por un conjunto de archivos que incluye formatos html, jpg, css y xml, en donde cada uno tiene una composición de cadenas de bits distintas y, además, es posible que el mismo libro cuente con ambas representaciones.

PREMIS puede aplicarse en cualquier repositorio interesado en la preservación digital, sin embargo, un factor clave en la implementación de los **metadatos de preservación** es que sus valores sean suministrados y procesados automáticamente por el ambiente de preservación o repositorio de preservación (PREMIS V3-LC, 2015), es decir que, estos datos

son los que se generan en una plataforma de preservación y que en su momento, pueden servir para trasladar de un ambiente a otro, los objetos, de ahí la importancia de su estandarización.

---

### 1.7.3. METADATOS PARA PRESERVACIÓN

Sin embargo, PREMIS no detalla elementos para la entidad intelectual, pues la descripción de una obra varía de acuerdo con el tipo de obra, de ahí la importancia de combinar éste esquema con otros esquemas de metadatos, por ejemplo, como Dublin Core, MARC 21, ONIX, MODS y METS que regularmente son utilizados para especificar libros digitales y determinar elementos de información *descriptiva* y en algunos casos *contextual, estructural, técnica y de derechos*, por lo que en un ambiente de preservación digital alimentan el proceso y se convierten en **metadatos para preservación**, ya que una vez dentro, forman parte de los metadatos de preservación al verse modificados por los acciones aplicadas para la conservación de los documentos y por lo tanto, contenidos dentro de un esquema PREMIS.

Así pues, ahora toca hablar de los **metadatos para preservación** que como ya se señaló, se refieren a la información descriptiva, contextual, estructural, técnica y en algunos casos de derechos y que de acuerdo con la norma ISO 23081-1:2006 es información requerida para incorporar un objeto a un sistema de gestión documental, con la posibilidad de mantenerlo a largo plazo.

Además, la misma norma señala que se trata de información relacionada con el contexto de creación, que incluye proceso de creación y origen del objeto además del contenido, apariencia, la estructura y atributos técnicos (REDC, 2008).

De ahí que para fines de esta investigación son metadatos que se van a asociar con lo que OAIS llama “productores” o creadores, además, estos metadatos encapsulan varios tipos de información, de ahí que existen diversos esquemas o estándares para su especificación, organización y control, pues de acuerdo con el tipo de objeto digital del que se trate, será el esquema que deba utilizarse, por lo que en esta investigación sólo se describirán esquemas de metadatos seleccionados por ser útiles en el ámbito de los libros.

---

#### 1.7.3.1. DUBLIN CORE

Dublin Core (DC)<sup>123</sup> es un esquema para información descriptiva de objetos, referente al contenido, la propiedad intelectual y la instancia del recurso, bajo el estándar de codificación XML<sup>124</sup> y de descripción RDF<sup>125</sup>, usado en internet por lo compacto y concreto en la estructura de información que maneja, basado en las siguientes 15 etiquetas (DCMI, 2017):

---

<sup>123</sup> Surge en 1995 y se formaliza en 1998 como estándar RFC 5791 y su última actualización fue con la ISO 15836-1:2017

<sup>124</sup> Extensible Markup Language.

<sup>125</sup> Resource Description Framework. Modelo de descripción de recursos que cuenta con una sintaxis de codificación con la cual es posible representar metadatos que pueden ser interoperables en servidores y clientes web desarrollados independientemente.

1. **Contributor.** Entidad responsable de hacer cambios en el recurso.
2. **Coverage.** Tiempo, espacio o jurisdicción en el cual el recurso es relevante.
3. **Creator.** Entidad responsable de la creación del recurso.
4. **Date.** Un punto o período de tiempo asociado con un evento en el ciclo de vida del recurso.
5. **Description.** Puede incluir un resumen, una tabla de contenido, una representación gráfica o relato en texto libre del recurso.
6. **Format.** El formato de archivo, el medio físico o las dimensiones del recurso.
7. **Identifier.** Una referencia inequívoca al recurso dentro de un contexto dado.
8. **Language.** Idioma del recurso.
9. **Publisher.** Una entidad responsable de hacer que el recurso esté disponible.
10. **Relation.** Un recurso relacionado, se recomendada colocar el URI del recurso con el que se relaciona.
11. **Rights.** Información de los derechos en y sobre el recurso, incluidos derechos de propiedad intelectual.
12. **Source.** Un recurso del que se deriva el recurso descrito, se recomendada colocar el URI del recurso del que se deriva.
13. **Subject.** Palabras clave, frases clave o códigos de clasificación que asignan una temática al recurso, se recomienda el uso de un lenguaje controlado.
14. **Title.** Nombre del recurso
15. **Type.** Sirve para asignar la naturaleza o el género del recurso, se recomienda usar un lenguaje controlado.

---

#### 1.7.3.2. MARC 21

El formato de codificación de información MARC<sup>126</sup> 21, es uno de los más usados y antiguo en el ámbito bibliotecológico para la catalogación de materiales documentales, es robusto en la cantidad de elementos que contempla para la descripción de estos y actualmente cuenta con una definición MARXML que permite el intercambio en plataformas abiertas con el protocolo OAI-PMH<sup>127</sup>. Se compone de tres elementos(MARC21-LC, 2008):

- **Cabecera.** Elemento de 24 caracteres, longitud fija que brinda información codificada respecto al registro.
- **Directorio.** Es una referencia a los campos variables dentro del registro y posee en 12 caracteres para cada etiqueta, la información del punto de inicio y longitud. Esta sección se cierra con un carácter ASCII 1EHex.
- **Campos variables.** Cada campo se identifica con una etiqueta numérica de tres caracteres y se clasifican en campos variables de control y campos variables de datos. Los primeros tienen una asignación numérica 00X, carecen de subcampos e indicadores y tienen una longitud fija con información codificada por posiciones. Los segundos, se agrupan en bloques de acuerdo con su etiqueta numérica, poseen dos

---

<sup>126</sup> Machine-Readable Catalog, diseñado como un esquema de información para registros bibliográficos propuesto en 1999 por la Biblioteca del Congreso de los Estados Unidos.

<sup>127</sup> Open Archive Initiative-Protocol for Metadata Harvesting

posiciones iniciales para indicadores y códigos de subcampos para marcar el inicio de cada elemento de información. El primer carácter de la etiqueta determina la función de los datos como se ve a continuación:

- 0XX: Información de control, números de identificación y clasificación, etc.
- 1XX: Asientos principales
- 2XX: Títulos y párrafo del título (título, edición, pie de imprenta)
- 3XX: Descripción física, etc.
- 4XX: Menciones de serie
- 5XX: Notas
- 6XX: Campos de acceso temático
- 7XX: Asientos secundarios diferentes a los de materias y series; campos ligados
- 8XX: Asientos secundarios de series, existencias, etc.
- 9XX: Reservados para implementación local

---

#### 1.7.3.3. ONIX

ONIX es un estándar internacional basado en XML para el intercambio de información descriptiva de libros, diseñado para la industria editorial desde editores, mayoristas, distribuidores, minoristas y colaboradores en la cadena de suministros del libro, la versión 49 de las listas de códigos u ONIX 3.0, es la vigente (EDItEUR, 2009).

Su objetivo principal es reducir los costos de inversión en sistemas y bases de datos para el control de las obras por cada uno de los participantes en la cadena de producción y distribución de un libro, de ahí que impulsa la estandarización de información para generar esquema de metadatos que acompaña al libro desde la creación y hasta la venta, evitando repetidos procesos de captura de información.

El esquema contiene información descriptiva del contenido de la obra, costo, medios de distribución e información que ayuda a la comercialización como críticas, premios, fragmentos y anexos importantes en cuanto a materiales complementarios. En Europa y Estados Unidos existen empresas dedicadas a dar servicios de captura, normalización y estandarización de datos en formato ONIX para la industria editorial, con posibilidades de que estos sean transformados a formato MARC 21.

---

#### 1.7.3.4. MODS Y METS

MODS<sup>128</sup> es un esquema XML para datos descriptivos particularmente diseñado para bibliotecas, permite transportar datos de registros MARC 21 o crear originales con etiquetas que son un subconjunto de campos MARC nombrados con lenguaje natural en lugar de números, sin embargo, este esquema no permite una transformación inversa y pasar los

---

<sup>128</sup> Metadata Object Description Schema. Creado por la oficina de Network Development and MARC Standards Office de la Biblioteca del congreso de Estados Unidos en 2002. Actualmente se encuentra en su versión 3.7 liberada en 2018.

datos de MODS a MARC 21, de ahí que existe la posibilidad de pérdida de información, por lo que resulta útil sólo en los siguientes casos (MODS-LC, 2020):

- En transferencia de registros MARC por cosechas en XML
- Como una extensión de METS para la parte descriptiva
- Para representar un registro MARC simplificado.

METS <sup>129</sup> permite contener metadatos descriptivos, estructurales y administrativos en lenguaje XML de libros producto de una digitalización. Es empleado para el intercambio de información entre repositorios digitales y consta de siete secciones (METS-LC, 2016):

1. **Cabecera:** Describe al propio documento METS.
2. **Metadatos descriptivos:** Puede contener los datos en las propias etiquetas METS o apuntar a esquemas externos como un registro MARC21 o EAD<sup>130</sup> almacenado en otro lugar e incluso ser una mezcla de datos contenidos o internos y ligas a externos.
3. **Metadatos Administrativos:** Se trata de información acerca de cómo se crearon y almacenaron los archivos que conforman el objeto digital, derechos de propiedad intelectual, datos sobre el objeto original a partir del cual se obtuvo la representación digital, e información sobre la procedencia de los archivos que conforman el objeto digital (es decir, relaciones entre copias maestras y derivadas, migraciones y transformaciones). Estos datos también pueden codificarse con el propio METS o ser externos.
4. **Sección Archivo:** Lista todos los archivos con contenidos que forman el objeto y los agrupa en las distintas versiones de este.
5. **Mapa Estructural:** Es la parte principal de un documento METS, contiene la estructura jerárquica del objeto digital y enlaza sus secciones con los archivos de contenido y los metadatos correspondientes a cada una de ellas.
6. **Enlaces Estructurales:** Permite registrar la existencia de hiperenlaces entre las secciones del mapa estructural, es muy útil cuando se usa METS para archivar sitios web.
7. **Comportamientos:** Se usa para vincular comportamientos ejecutables con los contenidos del documento METS. Cada comportamiento tiene una definición de interfaz y un "mecanismo" que identifica un módulo de código ejecutable que implementa y ejecuta el comportamiento definido de forma abstracta por la interfaz.

Todos estos esquemas de metadatos muestran una gran variedad de información que puede generarse acerca de un libro electrónico, sin embargo, depende del enfoque informativo en el que se analiza, lo que determina qué información es posible registrar del material, ya que mientras mayor sea el alcance de los metadatos para preservación, mayor

---

<sup>129</sup> Metadata Encoding and Transmission Standard, especificado por la Biblioteca del Congreso de los Estados Unidos en 2002

<sup>130</sup> Encoded Archival Description

puede ser la inversión de recursos en los procesos descriptivos, ya sea en las bibliotecas o con los creadores de los libros.

#### 1.7.4. LA PROBLEMÁTICA DE LOS METADATOS PARA PRESERVACIÓN

Como ya se mencionó, para la creación de los datos para preservación se pueden comenzar a crear desde la visión bibliográfica con MARC 21 o MARCXML y MODS, concentrada principalmente en la descripción del contenido intelectual aplicando reglas y normas de catalogación dentro de una biblioteca; o la visión de producción y distribución con ONIX, que da información generada por sus propios productores desde que se concibe la obra.

También están los datos que son generados principalmente para el intercambio y con sencillez con DC y METS, orientado al resguardo de objetos en repositorios digitales, que son generados por terceros y que en algunas ocasiones contienen aspectos estructurales, de derechos y administrativos cuando el libro se deriva de procesos de digitalización y que, además, permiten enlazar a esquemas descriptivos más amplios como MARCXML y MODS.

Sin embargo, las opciones para la descripción del *objeto* siguen siendo diversas, de ahí que resulta útil unificar las perspectivas y para eso el modelo conceptual de Requerimientos Funcionales para Registros Bibliográficos o FRBR<sup>131</sup> que permite analizar la composición del universo bibliográfico a partir de cuatro entidades conceptuales, más generales y aplicables a todos los esquemas de metadatos antes mencionados, ver Figura 11 (Tillett, 2004):

1. **Obra:** Representa al contenido intelectual.
2. **Expresión:** Se refiere al medio utilizado para la transmisión o comunicación de la obra, es decir lenguaje e idioma.
3. **Manifestación:** Se refiere a la materialización de la expresión, caracterizada por el formato o medio empleado para contenerla.
4. **Ítem:** Se refiere propiamente a la unidad física o virtual que va a representar a la manifestación y, por lo tanto, hace posible el acceso a la obra como una publicación perceptible y palpable con alguno de nuestros sentidos.

Figura 11. Entidades grupo1. Requerimientos Funcionales para Registros Bibliográficos



Fuente: Tillett, 2004.

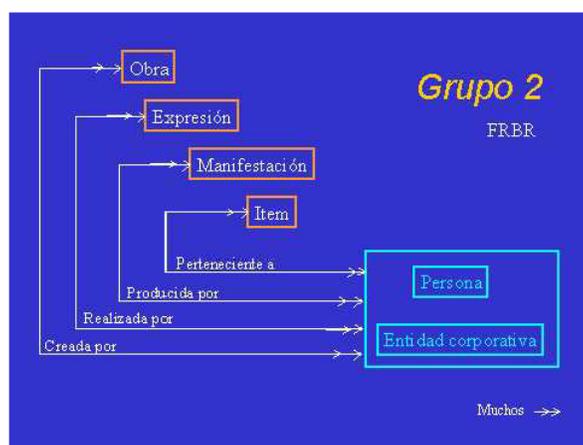
<sup>131</sup> Requerimientos Funcionales para Registros Bibliográficos (<https://www.loc.gov/catdir/cpsd/Que-es-FRBR.pdf>)

La forma de leer este diagrama del GRUPO 1 de FRBR es que una obra, puede realizarse en una o muchas expresiones, la expresión puede materializarse en una o muchas manifestaciones y ésta, se va a ejemplificar en uno o muchos ítems.

El modelo, también analiza con un diagrama entidad-relación<sup>132</sup> el tipo de vínculo que existe entre estas cuatro entidades conceptuales y entidades del mundo real, representadas como *persona / entidad corporativa* (Tillett, 2004).

5. **Persona / Entidad corporativa:** Se refiere al Individuo o grupo de personas que puede establecer una relación de creación, realización, producción o pertenencia con alguna de las entidades que conforman el universo bibliográfico.

Figura 12. Entidades grupo2. Requerimientos Funcionales para Registros Bibliográficos.



Fuente: Tillett, 2004.

La interpretación del diagrama del Grupo 2, ver Figura 12, puede hacerse de la siguiente manera: Una obra puede ser creada por una o más persona o una o más entidades corporativas, la expresión puede ser realizada por una o más personas o una o más entidades corporativas, la manifestación es producida por una o más personas o una o más entidades corporativas. En el caso del ítem, este puede pertenecer a una persona o una entidad corporativa, de ahí que la simbología utilizada en el diagrama puede ser incorrecta, pues debemos recordar que, al tratarse de una unidad física, no es posible que ésta pertenezca al mismo tiempo, a muchas personas o entidades corporativas.

Como ya se dijo, al comparar este modelo con los esquemas que se han revisado es posible ver que cada uno de estos, puede aplicar para describir alguna parte de las relaciones, pero no todas, por ejemplo, ONIX se centra en la relación de producción y MARC, describe lo necesario para una biblioteca, sin embargo esto puede resultar muy complejo para

<sup>132</sup> Lenguaje de modelado del ámbito de la Ingeniería de software conocido como modelo entidad relación o ERD, el cual ilustra relaciones y entidades que interactúan entre sí dentro de un sistema informático. En donde las entidades hacen referencia a personas, objetos o conceptos y las relaciones definen las conexiones entre ellas.

ambientes web en donde DC resulta útil, pero escaso cuando se trata de preservación digital, por lo que aquí entra otro esquema conocido como BIBFRAME<sup>133</sup>, que es una propuesta de descripción bibliográfica con un modelo de datos vinculados, con lo cual es posible aprovechar la riqueza de la información bibliográfica ampliando su utilidad más allá de las bibliotecas (BIBFRAME-LC, 2016) además de que considera el contenido conceptual y sus manifestaciones digitales.

BIBFRAME tiene por objetivo la transición del formato MARC 21 para facilitar el intercambio y reducir los costos de catalogación, para lo cual divide la información en tres niveles de abstracción o clasificaciones, ver Figura 11 (BIBFRAME-LC, 2016):

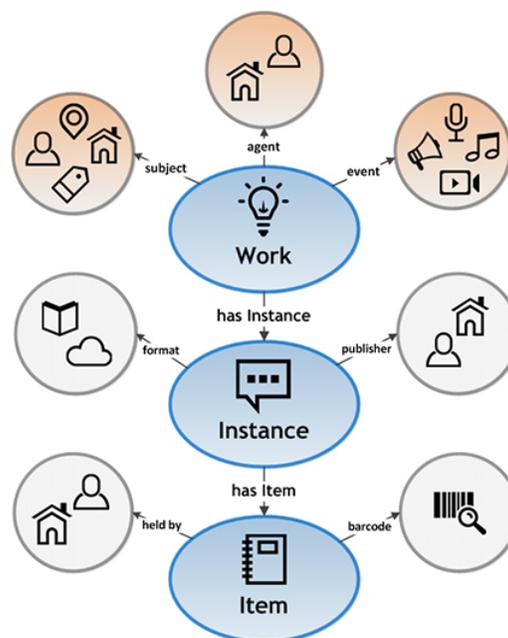
- **Work.** Se refiere a la Obra o esencia conceptual del recurso catalogado.
- **Instance.** Se refiere a las realizaciones materiales.
- **Item.** Se refiere al elemento físico o digital.

Y tres conceptos adicionales para la especificación del trabajo u obra:

- **Agents:** Se refiere a las personas o grupos de personas como organizaciones o entidades corporativas que se relacionan con un trabajo u obra e instancia, por ser autor, editor, ilustrador, etc.
- **Subjects:** Se refiere al sujeto o concepto del que trata el trabajo u obra. Un trabajo puede ser "sobre" uno o más conceptos. Este sujeto puede incluir conceptos, lugares, eventos, otras obras o instancias, personas, etc.
- **Events:** La forma en la que puede estar registrado el trabajo u obra.

Así pues, estos esquemas en su conjunto son alternativas que permiten darle una forma estandarizada a los metadatos en cuanto a su estructura, sin embargo, algo también muy importante es la manera en que estos se registran y para eso se han especificado normas que establecen códigos y pautas para crear los datos descriptivos de materiales bibliográficos que incluye a los libros.

Figura 13. Niveles de organización BIBFRAME 2.0.



Fuente: BIBFRAME, 2016.

<sup>133</sup> Bibliographic Framework. Desarrollado por la Biblioteca del Congreso de Estados Unidos en 2012 basado en clases y definiciones RDF.

RDA <sup>134</sup> es la nueva especificación para la codificación descriptiva de materiales bibliográficos, pensada para el entorno digital y ajustada para trabajar con el formato MARC y el modelo FRBR, que busca resaltar el alcance que tienen los datos catalográficos, por el valor que pueden proporcionar más allá de una biblioteca, de ahí que se basa como su nombre lo dice, en la descripción y el acceso (RDA, 2020).

Su intención es ser adaptable a las estructuras de datos para permitir la reutilización de los metadatos en diversas fuentes incluyendo la información que proporcionen los editores como ONIX y DC (Tillett, 2010). Pues la intención es contar con estructuras uniformes y contenidos de metadatos entendibles en cualquier ámbito.

Después de revisar estos esquemas y estándares, se tiene especificado, cuándo se generan los **metadatos de preservación digital** e incluso el cómo generarlos, pues como ya se dijo son los que se crean en el entorno de preservación, considerados también, parte de los metadatos administrativos.

Sin embargo, los **metadatos para preservación** están especificados sólo parcialmente pues el espectro de información que contienen es muy amplio, ya que va desde la descripción del contenido intelectual hasta la descripción del objeto digital y las propiedades o permisos que adquiere para los ambientes de consulta, además de que estos datos idealmente deben generarse desde que se crea el objeto.

También se tiene identificado que para la descripción del contenido intelectual de un libro digital, se tienen varios esquemas de metadatos, sin embargo su producción puede ser un proceso muy costoso dado que la información generada por los editores difiere de lo que requieren las bibliotecas, de ahí que iniciativas como RDA, FRBR y BIBFRAME, buscan apoyar la producción de estos metadatos descriptivos desde los editores, con la aplicación de estándares en ambos ámbitos para reducir los costos de catalogación en las bibliotecas.

Por otro lado, los metadatos que describe al objeto digital, conocidos en lo general como metadatos técnicos no tienen una claridad en su alcance, más allá de referencias en cuanto al formato de archivo, la cadena de bits y su entorno de ejecución, lo cual es vago, pues como veremos más adelante, esta información requiere una especificación mayor en cuanto a su alcance y que se orienta a su naturaleza como software.

El siguiente capítulo se centrará en el análisis de la edición del libro digital partiendo de identificar cuáles son los cambios que este genera en la cadena de valor del libro y así entender cuáles son los retos que los productores o editores del libro están enfrentando con la llegada del libro digital, para además comprender cuál es la información o metadatos técnicos que se generan derivados del proceso de producción de un libro desde la perspectiva de la ingeniería de software.

---

<sup>134</sup> Resource Description and Access.

## CAPÍTULO II. LA EDICIÓN DEL LIBRO DIGITAL.

El sector del libro atraviesa una crisis doble y profunda [...] una insoslayable crisis estructural que le llevará a ser algo distinto de lo que es, a operar de una manera nueva y diferente[...] donde la digitalización se constituye en una gran oportunidad o, a veces, en una gravísima amenaza para una reingeniería urgente de todo el ecosistema del libro (Gil & Rodríguez, 2011, p. 10)

En este capítulo se hará una revisión de los cambios que los medios digitales acarrearán a la edición del libro, como un eslabón en su cadena de valor o ciclo de vida a partir de establecer, cuál es su relación con la ingeniería de software desde la perspectiva del libro digital como un producto de software. También, se presentan los casos seleccionados e identificados para el análisis de la edición del libro digital, que son el punto de partida para el desarrollo del modelo que se presenta en el último capítulo de esta investigación.

### 2.1. LA EDICIÓN, CICLO DE VIDA Y CADENA DE VALOR DEL LIBRO

De acuerdo con Michael Bhaskar, editor de libros electrónicos, el término edición o en inglés publishing<sup>135</sup> puede referirse a una actividad y también a la industria, de ahí que el mismo Bhaskar menciona como “etimologías” (Bhaskar, 2014, p. 9) lo siguiente:

- 1.- La acción de dar a conocer algo públicamente; notificación oficial o pública; promulgación, anuncio público; 2.- La acción o empresa de preparar y poner en circulación libros, periódicos, etcétera, para venta o distribución públicas; (Bhaskar, 2014, p. 9)

Lo anterior basado en que, Bhaskar asegura que para algunos autores el matiz comercial de la palabra edición puede ser muy importante, entre los cuales cita como ejemplos los siguientes:

[...]la empresa o profesión de la producción y circulación comercial de literatura, información, partituras o en ocasiones grabaciones o arte. (Merriam Webster citado por Bhaskar, 2014, p. 32)

[...]la edición es un arte, un oficio y un negocio (Raymond Mortimer citado por Bhaskar, 2014, p. 14)

[...]la edición es el negocio de hacer cultura (Richard Nash citado por Bhaskar, 2014, p. 14)

[...]un negocio complejo que compra vende y fabrica o hace que se fabrique. Lo que compra y vende son productos de la imaginación humana, los materiales para hacer los libros y diversos derechos legales. Lo que fabrica nunca es lo mismo de un producto al siguiente (Diana Athill citada por Bhaskar, 2014, p. 14)

Además, respalda las dos etimologías mencionadas, por la relación que existe entre el concepto de edición y el de publicar o publish, en inglés, que de acuerdo con Michael Bhaskar, la palabra tiene para los libros un sentido más antiguo que la perspectiva industrial, pues refiere su existencia desde el año 1382 antes de la aparición de la imprenta, con un sentido más cercano a la declaración o anuncio de algo, por lo que publicar se refiere

<sup>135</sup> De acuerdo con Tomás Granados, las palabras no son equivalentes por completo, sin embargo, es posible emplearla.

también a “no ocultar” o “dar a conocer, anunciar o proclamar” que incluso incluye la lectura en voz alta dirigida a un público (Bhaskar, 2014, p. 30).

Bhaskar también remarca la importancia en diferenciar la edición del proceso creativo de la obra y el proceso de distribución o venta, por lo que es necesario diferenciar al editor, del autor o creador y el responsable de la venta. En el mismo sentido, Roger Chartier historiador y especialista en la historia del libro, menciona lo siguiente:

Es conveniente recordar que la producción de, no sólo de los libros, sino de los propios textos, es un proceso que, más allá del gesto de la escritura, implica diferentes momentos, diferentes técnicas, diferentes intervenciones: las de los copistas, los libreros editores, los maestros impresores, los cajistas, los correctores [...] El proceso de publicación, cualquiera que sea su modalidad, siempre es un proceso colectivo, que implica a numerosos actores [...] (Chartier, 2006)

Sin embargo, para Bhaskar la edición “nunca se separa del contenido” por lo que “una teoría de la edición surge de una teoría del contenido” (Bhaskar, 2014, p. 20) que parte de cuatro conceptos: marcos y modelos, filtrado y amplificación, pues:

El contenido se enmarca —se empaqueta para su distribución y se presenta a un público— de acuerdo con un modelo [...] Sin embargo, el verdadero núcleo de la edición reside en el filtrado y la amplificación. Publicar tiene que ver con seleccionar. Incluso quienes recurren a la autopublicación filtran: después de todo, eligen una obra —a saber, una hecha por ellos mismos— para publicarla. Aun en su versión más incluyente, la edición implica un proceso de filtrado marginal [...] El proceso completo de enmarcado en realidad está diseñado para amplificar textos. La edición tiene que ver con la expansión, a partir de un prototipo con el que se producen múltiples copias. El modelo es la razón por la cual se busca la expansión, por la cual se quiere amplificar [...] Si publicar significa algo, si el contenido público significa algo, es porque se apoya en la idea de la amplificación. (Bhaskar, 2014, p. 20)

Otra visión relacionada, es la de Robert Darnton, historiador estadounidense y especialista en la historia del libro, refiere que el libro impreso ha tenido a través de su historia un mismo ciclo de vida, descrito con sus palabras como:

un circuito de comunicaciones que va desde el autor hasta el editor (si el librero no asume ese rol), el impresor, el intermediario, el librero y el lector. (Robert Darnton editado por Finkelstein et al., 2006, p. 11)

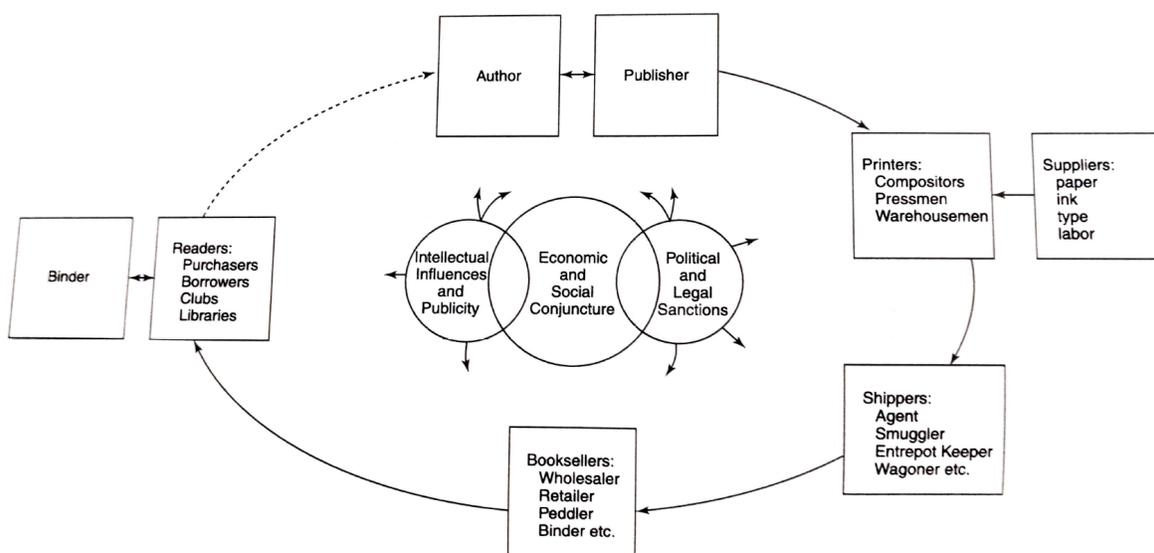
En el cual, de acuerdo con Darnton, el lector forma parte del circuito porque es una influencia para el autor, ya que este, dirige sus escritos a ciertos lectores tratando de anticiparse a sus reacciones y esto, junto con el hecho de que el autor también es lector, forman una retroalimentación continua.

De ahí que el libro, es un circuito para la comunicación de mensajes en el cual, el autor va a generar ideas para sus lectores, que deben pasar por un proceso cognitivo para llegar a la palabra escrita (editores e impresores) además de involucrar a otros participantes (libreros,

transportistas, distribuidores, etc.) que intervienen para la construcción de una publicación y que permiten su llegada con el lector (dentro de los cuales se incluye a las bibliotecas) con lo cual se establece un ciclo, pues el lector tiene la posibilidad de retroalimentar al autor y así, este último puede enriquecer los constructos que genere en sus siguientes obras e incluso modificar lo que ya ha producido.

Pero además este circuito de comunicación del libro se ve influenciado por factores sociales, económicos, políticos, legales, de mercado e intelectuales del momento en que se concibe el libro, afectando a todos los participantes en el circuito, ver Figura 14 (Robert Darnton editado por Finkelstein et al., 2006, p. 11).

Figura 14. Circuito de comunicación del libro.



Fuente: Robert Darnton editado por Finkelstein et al., 2006, p. 11.

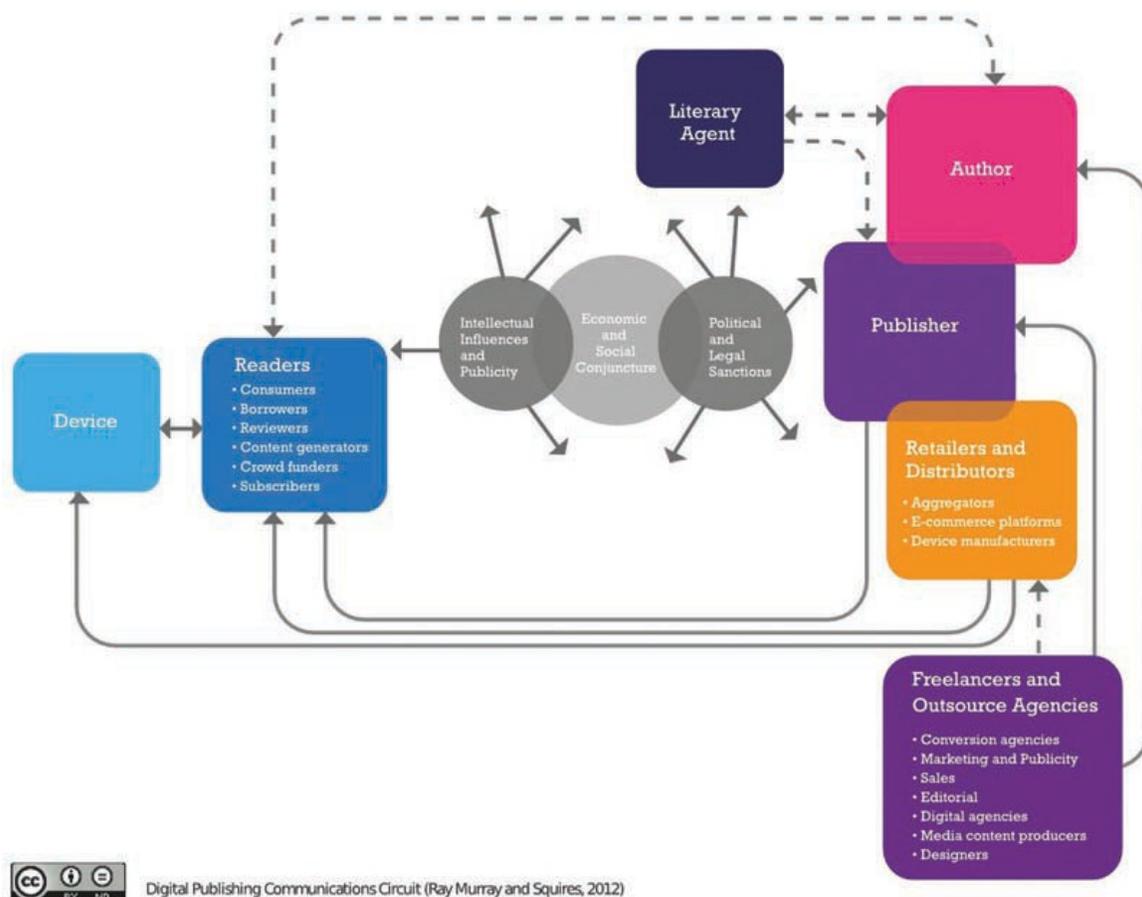
Estas perspectivas muestran que, tanto el contenido como el proceso de comunicación son fundamentales para la conformación del libro y ambos elementos confluyen en la edición que incluye la actividad comercial e implica la inversión o utilización de recursos de diversa índole y el hecho, de dar a conocer el contenido o distribuirlo para que llegue a un lector dentro de los cuales se incluye a las bibliotecas.

Esto último, nos lleva recordar que en esta investigación ya se habló de las bibliotecas como lugares destinados para que los lectores tengan acceso a los libros, por lo que además con la propuesta de Darnton se reafirma este papel, pero además, también ya se dijo que existen archivos y bibliotecas responsables del resguardo y preservación del libro, las cuales regularmente cuentan con políticas y legislaciones nacionales que apoyan su labores, como el Depósito Legal, elementos que también se reflejan en el circuito del libro definido por Darnton.

Además en este circuito de comunicación Darnton se refiere la influencia que se crea en el libro a partir de la relación que existe entre el autor, el editor y el lector, que como ya se dijo incluye a la biblioteca, lo cual también se refleja como factor determinante en la preservación digital, pues el modelo OAIS lo menciona como “La negociación con quienes producen los objetos de información, para el ingreso apropiado al Archivo” (CCSDS, 2012, pp. 3-1) que como ya se vio en el capítulo anterior, contempla la integración de metadatos.

De ahí que el ciclo de vida planteado por Darnton, también concentra en su esencia estos elementos como factores inmersos en la vida del libro digital, sin embargo, Murray & Squires, señalan que son “cambios sísmicos” los que transforman el circuito de Darnton en el ámbito digital, pues “los autores ahora pueden publicar, comercializar y distribuir su trabajo sin la ayuda de un editor” (Murray & Squires, 2013, p. 4), ver Figura 15.

Figura 15. Digital Publishing Communications Circuit.



Fuente: Murray & Squires, 2013, p. 8.

Por lo que, Murray & Squires, plantean un nuevo circuito para el libro digital, en el cual la figura del autor se sobre poner a la del editor y, además, los lectores se convierten “en creadores de contenido” con lo cual se genera un “un nuevo ecosistema editorial”, pues se hace más evidente la relación entre el lector y los autores. También, en este nuevo esquema

las bibliotecas se ven afectadas pues se convierten en espacios virtuales que requieren de un dispositivo de lectura y en todo el circuito, surge la presencia de un nuevo intermediario, especialistas en medios digitales, necesario como apoyo para los autores, editores y distribuidores y que como se verá más adelante, para fines de esta investigación, es lo que genera una relación con la Ingeniería de software y la edición del libro (Murray & Squires, 2013, p. 4).

Así pues, los cambios en la comercialización y producción o edición tradicional son evidentes en el libro digital de ahí que otros especialistas como Bhaskar, también lo menciona como un cambio en la *cadena de valor*, término que el autor no define, sin embargo, se identificó que se trata de un concepto utilizado en el ámbito de la economía para establecer:

las principales actividades que crean un valor para los clientes y las actividades de apoyo relacionadas. La cadena permite también identificar los distintos costos en que incurre una organización a través de las distintas actividades que conforman su proceso productivo, por lo que constituye un elemento indispensable para determinar la estructura de costos de una compañía. Cada actividad en la cadena de valor incurre en costos y limita activos, para lograr su debido análisis y consideración permiten mejorar la eficiencia tecno-económica de una empresa, un grupo de empresas o de un determinado sector industrial. (Quintero, 2006)

En este sentido, se encontró que este término aplicado en el ámbito del libro se define como la *cadena de valor del libro* que, de acuerdo con Ernesto Piedras, economista y estudioso de la industria editorial en México, la describe como:

La cadena de valor del libro puede ser dividida en dos grandes niveles: el de creación y fabricación, y el de distribución. En estos eslabones participan tres importantes personajes, que son los autores, los editores y finalmente los lectores. Sin embargo, es importante mencionar que la creatividad de los autores representa el insumo esencial que detona la cadena de valor de la Industria Editorial. (Piedras, 2013, p. 35)

Además, Piedras en un estudio que elaboró en 2013 respecto a la industria editorial en México, refiere algunos conceptos establecidos en la *Ley de Fomento para la Lectura y el Libro*, mexicana:

**Edición:** Proceso de formación del libro a partir de la selección de textos y otros contenidos para ofrecerlo después de su producción al lector (Piedras, 2013, p. 22)

**Editor:** Persona física o moral que selecciona o concibe una edición y realiza por sí o a través de terceros su elaboración (Piedras, 2013, p. 22)

**Cadena productiva del libro:** Conjunto de industrias que participan en los diversos procesos de producción del libro y está conformada por la de la Celulosa y el Papel, la de las Artes Gráficas y la Editorial. En la de las Artes Gráficas se incluye la participación de los que brindan servicios editoriales, los impresores y los encuadernadores que reciban sus ingresos en más de un ochenta por ciento de los trabajos relacionados con el libro y la revista (Piedras, 2013, pp. 22-23)

**Autor:** Persona que realiza alguna obra destinada a ser difundida en forma de libro. Se considera como autor, sin perjuicio de los requisitos establecidos en la legislación vigente, al traductor respecto de su traducción, al compilador y a quien extracta o adapta obras originales, así como al ilustrador y al fotógrafo, respecto de sus correspondientes trabajos (Piedras, 2013, p. 23)

Es importante señalar que para los fines de esta investigación estos conceptos son de utilidad, a pesar de que en otros ámbitos o desde otras perspectivas de estudio, pudieran ser tema de amplias discusiones y análisis, sin embargo, este trabajo no pretende hacer un estudio profundo respecto a la teoría de la edición por lo que, estas concepciones se presentan únicamente para generar un contexto que sirva de referencia para el análisis que se expondrá más adelante.

Volviendo al tema de los cambios en la edición, Robert Chartier lo señala como una “revolución digital” que para el libro representa una “revolución de la modalidad técnica” de la producción y circulación, además de una “revolución en la percepción” y “prácticas de lectura” pues:

obliga al lector contemporáneo a abandonar todas las herencias pasadas [...] ya que la textualidad digital no utiliza más la imprenta, por lo menos en su forma tipográfica, ignora el libro unitario y está ajena a la materialidad del código [...] También está el desafío de identificación del libro entendido como obra, el libro entendido como objeto cuya existencia empezó durante los primeros siglos de la era cristiana y que parece desaparecer en el mundo de los textos electrónicos (Chartier citado por Rosales, 2012).

En ese mismo sentido, Manuel Gil y Joaquín Rodríguez, reconocidos editores españoles, refieren que el futuro del libro ahora se encuentra inmerso en “un proceso de mutación digital” que impacta directamente en la cadena de valor del libro, en el cual es necesario decidir cuáles son las competencias, conocimientos y herramientas que se deben adquirir para que los editores, libreros, distribuidores y bibliotecarios puedan hacer una transición del proceso tradicional de edición al digital de forma ordenada y rigurosa, en beneficio de su “cliente final”, el lector (Gil & Rodríguez, 2011, p. 9).

Además, Gil y Rodríguez refieren que se trata de un “nuevo algoritmo de producción, distribución y consumo abierto” (Gil & Rodríguez, 2011, p. 34) sustentado en Internet con una arquitectura de diez ejes (Gil & Rodríguez, 2011, pp. 34-44):

1. Un paradigma de acceso abierto
2. Canal de distribución y marketing
3. Posibilitador del proceso de desintermediación
4. Irruptor de nuevos modelos de negocio
5. Tsunami de producción y generación de contenidos incluida la Web 2.0
6. Ocaso de los monopolios de distribución
7. Apoyo al formato papel

8. Facilitador de la irrupción de nuevas propuestas de valor, pues cada usuario se convierte en generador de contenidos
9. Medio que obliga a las editoriales y librerías a repensar su naturaleza.
10. Marca y comunidad con valores simbólicos tangibles e intangibles, pues el editor juega un nuevo rol como creador de seguidores e integrador de comunidades en el ciber espacio

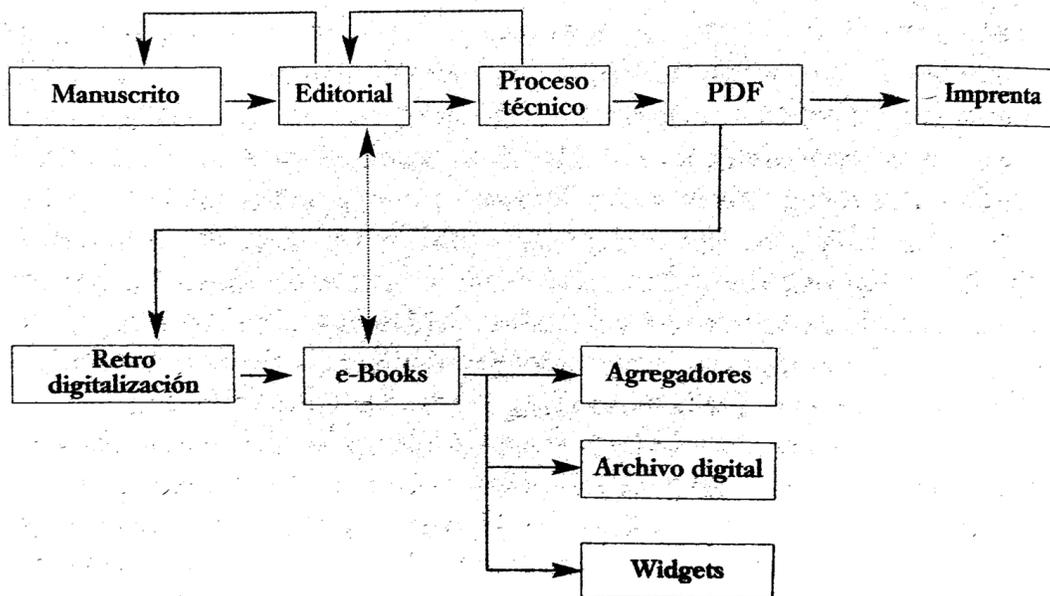
Al respecto, Manuel Gil y Joaquín Rodríguez, plantean que el editor requiere asimilar la transformación de la industria editorial a una “empresa red” y/o “empresa comunidad” (Gil & Rodríguez, 2011, pp. 29-33) que implican la colaboración entre editores y con otras industrias, como las de tecnología y agentes ajenos al campo editorial o en su defecto grandes inversiones en infraestructura y mano de obra calificada para el manejo de tecnología. Además, los autores señalan el inminente daño que la transformación digital y la falta de adaptación conllevan, pues de facto, ha comenzado a generar la desaparición de imprentas, encuadernadoras, distribuidoras tradicionales y librerías, estas últimas con una urgente necesidad de reconversión, pues compañías como Amazon o iBookStore se han vuelto fuertes competidoras. De ahí que Gil y Rodríguez establecen 7 principios estratégicos como tesis para la supervivencia del editor en la era digital (Gil & Rodríguez, 2011, pp. 91-92):

1. *Tesis de la Convergencia*: El eje fundamental del editor debe ser el contenido, sin importar el formato y soporte.
2. *Tesis de la Integración*: El editor debe considerar una constante generación de valor adicional para sus usuarios o lectores.
3. *Tesis de la Competencia tecnológica*: Las habilidades tecnológicas son una ventaja, por lo que deberán dominar aquello que afecte su negocio y contar con colaboradores o partners en tecnología. Un error en este sentido genera consecuencias funestas.
4. *Tesis 1:1 (One to One)*: La hiperfragmentación de contenidos conlleva fuertes personalizaciones en el consumo.
5. *Tesis A3 (Anything, Anywhere, Anytime)*: Las nuevas tecnologías están encaminadas a recibir cualquier clase de contenido, en cualquier lugar y en cualquier momento, por lo que es necesario generar una oferta editorial dinámica.
6. *Tesis de retorno publicitario*: Es necesario contemplar internet como un medio publicitario y las editoriales pueden integrarlo como un mecanismo más de ingresos.
7. *Tesis de comprensión*: La digitalización conlleva un rediseño de la cadena de valor el cual no es un proceso de negociación, es posicionamiento competitivo.

Gil y Rodríguez, esquematizan lo que llaman “digitalización estandarizada” (Gil & Rodríguez, 2011, p. 88) al hablar del impacto de lo digital en la edición, en la cual plantean un flujo que comienza con el manuscrito y termina en la imprenta, pero además incluye un proceso

técnico que interactúa con el editorial y que genera un PDF, lo cual se va a la imprenta, sin embargo, también del PDF se genera en un e-book que se deriva en un archivo digital, ver Figura 16.

Figura 16. Digitalización Estandarizada.



Fuente: Gil & Rodríguez, 2011, p. 88.

Como puede verse la transformación en la cadena de valor o producción del libro que afecta a la edición involucra la participación de especialistas en tecnologías o en su caso como lo señala la *tesis de la competencia tecnológica* de Gil y Rodríguez, requiere dominar ciertas habilidades tecnológicas que ayuden al editor a comprender esta nueva forma de producción, creando un nuevo perfil, el editor digital.

Sin embargo, esto podría parecer contradictorio pues también existe el concepto de autoedición como un cambio importante impuesto por la era digital, que Murray & Squires mencionan como determinantes y disruptivos, pero Bhaskar los señala sólo como la posibilidad que de que “todo el mundo cuente con las herramientas para crear y divulgar contenido con oprimir un botón” (Bhaskar & Rubio Ruiz, 2017), y que desde la opinión del mismo Bhaskar, no es algo que afecte a la edición.

Sin embargo, Silvia Bueno Senz en su artículo «En un lugar de la “Mancha”...» *Procesos de control de calidad del texto, libros de estilo y políticas editoriales*, refiere una gran preocupación por la rapidez con la que la tecnología permite crear un libro, por lo que hace un llamado a no olvidar el propósito ancestral que tiene la edición de un libro y recalca la importancia de continuar con el cuidado de la calidad del contenido, en lo que ella señala como edición tradicional (Bueno, 2005).

De ahí que Silvia Bueno, recalca que los procesos de control de calidad del texto deben ser los mismos a pesar de los avances en las herramientas de composición e impresión, pues estas sólo han permitido simplificar o acelerar algunos pasos. De ahí que los procesos para controlar la calidad de un texto desde su perspectiva, siguen siendo (Bueno, 2005):

1. Auxiliar a un autor que presenta carencias estilísticas más o menos relevantes.
2. En el caso de las traducciones, dar al lector garantías de integridad y fidelidad en la traslación de un texto de un idioma a otro.
3. Adecuar un texto a una serie de convenciones que persiguen facilitar su lectura y comprensión al lector a quien va destinado.
4. Conferir al texto un nivel de corrección y depuración lingüística que lo haga más eficaz como vehículo de un mensaje.
5. Dar al texto, mediante la aplicación de pautas estilísticas propias de cada editorial, un sello específico que lo distinga de otros.
6. Generar con todo ello una imagen de prestigio y calidad ante el lector, de la que se beneficiarán tanto el propio autor como la casa editorial que hace público su trabajo.

Cabe resaltar que en esta investigación no se pretende ahondar en lo que se refiere a la calidad de los textos de un documento o libro, pues es un tema que requiere un estudio que sale de los alcances de este trabajo, sin embargo, vale resaltar que el cuidado de los textos en el ámbito digital debe apegarse a lo que tradicionalmente se maneja en el proceso de edición, a pesar de los beneficios de la autoedición y las nuevas dinámicas de escritura que tampoco son tema de esta investigación.

De ahí que, lo que sí es importante resaltar con todo esto, es que desde la perspectiva del libro digital y sus características como software, Gil y Rodríguez plantean la existencia de un editor en el ámbito digital, lo cual, además, fundamenta el desarrollo de nuevas habilidades que debe adquirir el editor para la producción de libros digitales.

Así pues, ahora es necesario hablar de la esencia de la Ingeniería de Software y la importancia de ésta, derivada de la visión establecida en este trabajo de investigación del libro digital como software, pues más adelante, esto permitirá identificar los elementos que hablan de la calidad del libro en el ámbito digital y del papel que debe jugar el editor digital, ya que hasta el momento sólo se han mencionado, pero no se ha especificado.

## **2.2. LA ESENCIA DE LA INGENIERÍA DE SOFTWARE APLICABLE AL LIBRO.**

Para hablar de la aplicación, importancia y la relación entre la Ingeniería de Software y la edición de los libros digitales, conviene comenzar por entender cuál es el origen de la Ingeniería de Software para comprender el ámbito y alcance de esta.

En cuanto al uso del término Ingeniería de Software, existen dos momentos importantes en su historia, primero el origen del término y luego la formalización de éste en el área de la computación. Respecto a su origen se tienen identificados dos instantes, el primero con

Margaret H. Hamilton<sup>136</sup> (Cameron, 2018) (H. J. Rubio, 2014) (Rayl, 2008), encargada de la construcción del software para el Apollo en 1965, quien se dice, empleó el término para legitimar y colocar en igualdad de importancia el trabajo de su equipo respecto al resto de los ingenieros de la NASA, sin embargo existe otra versión, ya que hay quienes refieren que el término lo empleo por primera vez Anthony A. Oettinger<sup>137</sup> en 1966 (Meyer, 2013) al escribir una nota editorial en la revista ACM para referirse a la importancia de los profesionistas relacionados con la informática y la computación, encargados de la construcción de software.

Sin embargo, respecto a la formalización y oficialización del término, esto se dio en 1968 durante una reunión organizada por la OTAN en la que participaron profesionales de once países quienes discutieron aspectos que incluyeron el diseño, la producción y la relación del software con el hardware, además de su distribución, comercialización y uso como servicio (Randell, 1996) (Naur & Randell, 1969).

Esta convocatoria de la OTAN tuvo su origen a partir de un suceso conocido en el ámbito de la computación como “crisis de software” (Dijkstra, 1972) pues en los años 60’s se produjo un problema importante en el ámbito del cómputo industrial que llevó a una crisis, debido a la falta de personal capacitado en la construcción de software, de ahí que los países miembros de la OTAN convocaron a especialistas quienes discutieron la importancia en garantizar la confiabilidad y producción de software, dado el impacto que éste comenzó a tener en actividades centrales de la sociedad (Naur & Randell, 1969). Algunas reflexiones hechas por los asistentes en la reunión de 1968 fueron:

Berghuis<sup>138</sup>: No existen los paquetes de software independientes; estos se ejecutan en un equipo (hardware), necesitan procedimientos para ser operados...También debemos considerar su organización desde el punto de vista del desarrollo de sistemas y, de hecho, enfrentar las diferencias entre la organización funcional y la del proyecto.<sup>139</sup> (Naur & Randell, 1969, p. 15).

David<sup>140</sup> and Fraser<sup>141</sup>: ...un desafortunado método de desarrollo y producción de un proyecto... conduce a un producto con gran cantidad de errores, inflexible y difícil de manejar. Es poco probable que un producto así sea confiable o que pueda mantenerse ...<sup>142</sup> (Naur & Randell, 1969, p. 71).

---

<sup>136</sup> Margaret H. Hamilton es CEO de Hamilton Technologies, Inc. Es matemática del Instituto de Tecnología de Massachusetts, creó lo que ella llamó software ultra confiable para el viaje a la luna. Se dice que para resaltar la importancia de su trabajo y legitimar a su equipo frente al resto de los ingenieros de la NASA ocupó el término, equiparando la ingeniería mecánica y en computación con la ingeniería de software.

<sup>137</sup> El Dr. Oettinger, lingüista e informático, pionero en el desarrollo de códigos de computadora e inteligencia artificial (IA) y escribió los primeros programas de IA para incorporar el aprendizaje.

<sup>138</sup> Dr. J. Berghuis, director de Philips’ Computer Industrie (Los países bajos)

<sup>139</sup> Traducido por Ana Yuri

<sup>140</sup> Dr. E.E. David, Jr., director ejecutivo de la división de sistemas de comunicaciones de Bell Telephone Laboratories. (USA)

<sup>141</sup> Mr. A.G. Fraser, miembro del laboratorio de matemáticas, Cambridge, (Inglaterra)

<sup>142</sup> Traducido por Ana Yuri

Babcock<sup>143</sup>: ... es responsabilidad del fabricante generar sistemas a la medida para satisfacer las necesidades del usuario ... En la verificación previa al lanzamiento, está el cómo un fabricante puede garantizar que un sistema funcione adecuadamente en un entorno conocido.<sup>144</sup> (Naur & Randell, 1969, p. 66).

Gillette<sup>145</sup>: ... el mantenimiento de sistemas es un proceso interminable que dura toda la vida útil ... En resumen, entonces, el proceso de mantenimiento implica corrección de código, mejora de código y extensión de código.<sup>146</sup>(Naur & Randell, 1969, p. 65).

Selig<sup>147</sup>: Con la rápida proliferación de lenguajes informáticos, subrutinas y programas, y el tremendo esfuerzo que representan, la documentación meticulosa es esencial, no solo para ahorrar dinero además para prevenir el caos.<sup>148</sup> (Naur & Randell, 1969, p. 68).

Como puede verse en los ejemplos expuestos respecto a los comentarios vertidos por los especialistas en 1968, los problemas con la construcción del software pueden trasladarse a todos los ámbitos en los que el software se utiliza y produce, por lo que no sería extraño que algún comentario expresado en esa época pudiera resultar familiar en el ámbito editorial, aún en este momento, de ahí la importancia de entender que la construcción del software visto sin una perspectiva formal genera problemas de forma natural, provocando una crisis en los productos por la falta de aplicación de elementos básicos para su construcción.

Así pues, la Ingeniería de Software, se ha definido en el ámbito de la computación como un área encargada del estudio y aplicación de prácticas para la construcción del software por lo que, de acuerdo con lo establecido en su base de conocimientos, el estándar ISO/IEC TR 19759:2016 conocido como SWEBOK<sup>149</sup> (Bourque et al., 2014), se define como:

la aplicación de un enfoque sistemático, disciplinado y cuantificable al desarrollo, operación y mantenimiento de software. (SEVOCAB et al., 2021e)

En donde, el desarrollo de software sea cual sea el método empleado, contempla seis elementos básicos (Bourque et al., 2014, pp. 8-3) ver Figura 17:

1. se alimenta de entradas,
2. un conjunto de criterios que deben identificarse y serán aplicados a dichas entradas,
3. actividades de transformación que hacen uso de las entradas y criterios,
4. que a su vez generan tareas,
5. las cuales después de cumplir con criterios de salida,

---

<sup>143</sup> Mr. J.D. Babcock, fundador de la empresa Allen-Babcock Computing (USA)

<sup>144</sup> Traducido por Ana Yuri

<sup>145</sup> Mr. H.R. Gillette, miembro de Control Data Corporation, (USA)

<sup>146</sup> Traducido por Ana Yuri

<sup>147</sup> Dr. F. Selig, miembro de Mobil Research and Development Corporation (USA)

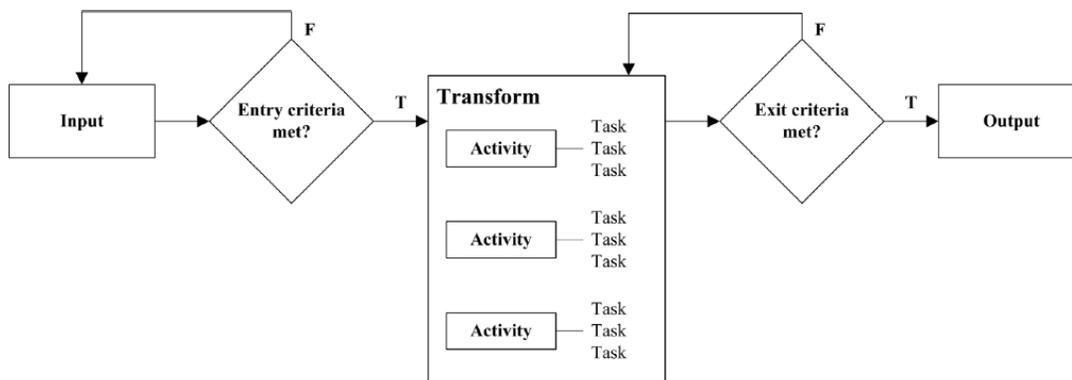
<sup>148</sup> Traducido por Ana Yuri

<sup>149</sup> Software Engineering Body of Knowledge. La Guía del Cuerpo de Conocimientos de Ingeniería de Software o SWEBOK publicada en 2004 por la IEEE. Busca orientar respecto al conjunto de conocimientos desarrollados durante las últimas cuatro décadas, para promover el avance tanto de la teoría como de la práctica de la profesión.

6. generan un resultado, conocido como software.

Además, cabe resaltar que estos elementos deben contemplar el establecimiento de acciones generales que ayuden a controlar el proceso ejecutado para la creación de software, de tal forma que se facilite la comunicación y comprensión de las necesidades humanas que surgen de esta actividad y así mejorar la calidad de los productos generados (Bourque et al., 2014).

Figura 17. Elementos del proceso de software.



Fuente: Bourque et al., 2014, pp. 8-3.

Esta forma abstracta de conceptualizar la construcción de software, al compararla lo presentado por Gil y Rodríguez en la sección anterior, respecto a la digitalización estandarizada de la edición, puede compararse e igualarse si consideramos el manuscrito como la entrada en el proceso de construcción del libro digital, en donde, el PDF o e-book es la salida y las actividades o tareas de transformación se concentran en la edición y los procesos técnicos ver Figura 16 y Figura 17.

Además, esos elemento o criterios de entrada y salida señalados por Bourque, también concuerdan con el concepto de filtrado expuesto por Michael Bhaskar, en donde explica que todo contenido requiere cierto proceso de selección que ayuda a perfilar lo que se quiere ofrecer al lector. Y que desde la perspectiva de la Ingeniería de software se conoce como necesidades de usuario, de las cuales se hablará más adelante.

Así pues, retomando la visión de lo que es la Ingeniería de Software y la visión definida en el área, respecto a los procesos involucrados en el desarrollo, operación y mantenimiento de software, en 2013 se publicó por el SEMAT<sup>150</sup>, un núcleo o kernel de elementos esenciales, conocido como ESSENCE, en el cual se describen las "cosas con las que siempre se trabaja", "cosas que siempre se hacen" y "habilidades que es necesario tener" (Jacobson et al., 2012) para todos los esfuerzos relacionados con el software, el cual provee las siguientes herramientas (Jacobson et al., 2013):

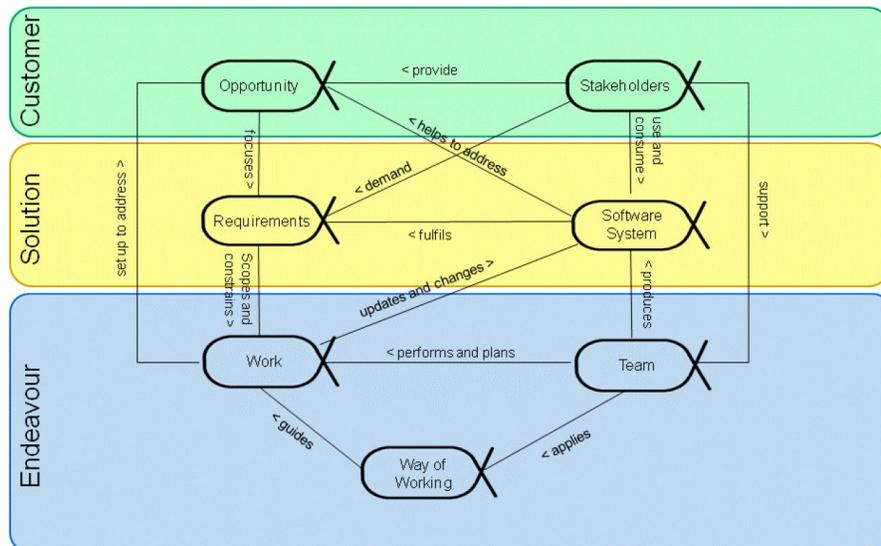
<sup>150</sup> SEMAT (Software Engineering Methods and Theory). fue fundada en 2009 por Ivar Jacobson, Bertrand Meyer y Richard Soley.

- Un marco de pensamiento para que los equipos razonen sobre el progreso de su trabajo y la salud de sus esfuerzos.
- Un marco para que los equipos ensamblen y mejoren continuamente su forma de trabajo.
- El terreno común para la comunicación mejorada, la medición estandarizada y las mejores prácticas compartidas.
- La base para las definiciones de métodos y prácticas interoperables y accesibles.
- Y una forma de ayudar a los equipos a comprender en dónde se encuentran y qué deberían hacer.

El núcleo se divide en tres áreas: *cliente*, *solución* y *esfuerzo* y dentro de estas existen elementos esenciales de *cosas con las que se trabaja*, *cosas por hacer* y *habilidades necesarias*. Además, una característica importante de ESSENCE es la posibilidad que otorga de que cada equipo defina los métodos, procesos o prácticas que mejor se adecuen al desarrollo de un software con base en el equipo, recursos y producto del que se trate.

Los primeros elementos, llamados *alpha*<sup>151</sup> por sus siglas en inglés, *se refieren a las cosas con las que se trabaja* y permiten evaluar en cualquier momento y de acuerdo con cada área del núcleo de ESSENCE, el progreso y salud del proceso de software que se lleva a cabo, ver Figura 18.

Figura 18. “The Kernel Alphas” o alphas del núcleo.



Fuente: OMG, 2015, pág. 17.

Cada *alpha* se compone de un conjunto de *estados alpha*, los cuales, disponen de listas de verificación que funcionan como criterios para validar continuamente el progreso y saber, si se tiene un verdadero avance en el desarrollo y si el método seleccionado es el adecuado al

<sup>151</sup> Abstract-Level Progress Health Attribute

desarrollo, basado en las expectativas generales de un producto de software. Esto puede ser útil durante la edición de un libro digital, pues los editores podrían seleccionar los procesos que deseen emplear para la creación del libro con base en los alcances y necesidades que se quieran satisfacer en el lector, por lo que el proceso empleado puede ser tan complejo o sencillo como sea el tipo de libro electrónico del que se trate.

Dentro del área *cliente* existen dos *alpha*: *oportunidad e interesados*, la primera se centra en el análisis de la motivación o circunstancias que conducen a la creación o modificación de un software. De ahí que los *estados* de esta *alpha* contemplan elementos que ayudan al equipo destinado al trabajo, a confirmar que comprenden y comparten las necesidades de los interesados en el software o *dominio del problema*<sup>152</sup> (OMG, 2015). Esta visión de ESSENCE puede dirigirse a las necesidades que el libro busca satisfacer respecto a su lector, pero además se relaciona directamente con las nuevas formas de lectura en el ámbito digital, que pueden generarse en caso de que el editor busque innovar o crear nuevas oportunidades para su usuario o si es el caso, sólo considerar las nuevas exigencias de los lectores.

El *alpha interesados* se concentra en las personas u organizaciones que intervienen, se relacionan o son afectados por el software en cuestión, pues estos son los que generan los requisitos a satisfacer, aportan los recursos para hacerlo y se ven afectados por las oportunidades o beneficios que del software se deriven. Por lo tanto, todos los estados que componen a esta *alpha* se orientan para confirmar que las partes interesadas trabajan coordinadamente con el equipo en el proceso que se lleve a cabo (OMG, 2015).

Para el caso del libro, los interesados son todos aquellos que intervienen en su cadena de valor y ciclo de vida que como Darnton menciona, van desde el autor, editor, impresor, librero, distribuidor y lectores, en donde también se incluyen a las bibliotecas. Esta visión es muy importante pues parte esencial en la construcción del software es que este considere elementos para satisfacer las necesidades que cada interesado requiere.

En el área *solución* se encuentran dos *alphas*: *requisitos y sistema de software*. Los requisitos se refieren a lo que el software debe hacer para generar una oportunidad que satisfaga a las partes interesadas. Y el sistema de software se refiere al proceso de composición de software, hardware y datos que al trabajar coordinadamente generan el valor para las partes interesadas, de ahí que contempla un proceso que considera la comprensión de los alcances del software, el diseño de la arquitectura en la que interactúan software, hardware y datos, y la codificación del software, además de incluir la comprobación de que el producto cumple con su objetivo (OMG, 2015).

Esta *alpha*, proporciona para el nuevo proceso de edición del libro la integración de actividades técnicas necesarias y derivadas de lo que implica crear un software, que son: la

---

<sup>152</sup> Se refiere al ámbito en el que se desenvuelve el software, es decir se refiere a conocer cuáles son las necesidades que se deben satisfacer de acuerdo con el tipo de organización, industria, público o usuario para el que se desarrolla el producto.

comprensión del alcance del libro, el diseño de su arquitectura basada en el hardware y datos o metadatos que lo conforman, la codificación de la obra como libro , también referida como arquitectura de la información y la comprobación de que el libro cumple con las necesidades o alcances planteados para la edición, también conocida como calidad del software.

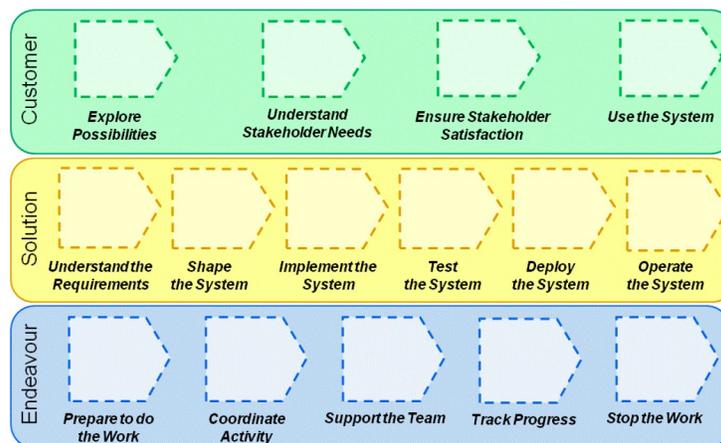
La última área, *esfuerzo*, contempla tres *alphas*: *trabajo*, *equipo* y *forma de trabajar*, en donde la primera se refiere al esfuerzo mental o físico que debe realizarse para alcanzar el resultado deseado, el *equipo*, contempla a las personas que participaran activamente en el *trabajo* y la *forma de trabajar*, es el conjunto de prácticas y herramientas utilizadas por el *equipo* para realizar el *trabajo*.

Los estados que componen cada una de estas *alphas* permiten monitorear la existencia de estas tres instancias, así como su efectividad en lo general, de tal forma que “el equipo evoluciona su forma de trabajar junto con su comprensión de su misión y su entorno de trabajo” con forme avanza en el desarrollo del software y “reflexionan continuamente sobre su forma de trabajar” para adaptarla “según sea necesario”(OMG, 2015, p. 18).

Para la edición del libro digital, esta última área de ESSENCE permitiría al editor comprender cómo hacer la composición de su equipo de trabajo y establecer nuevas formas para la edición que se adapten a las necesidades de cada obra que desee publicar, derivado de las nuevas necesidades del libro en el ámbito digital.

En cuanto a las cosas por hacer, el núcleo tiene definido para cada espacio de actividad elementos que permitirán guiar al equipo de forma general en el proceso, sin importar qué prácticas o métodos se emplean, ver Figura 19.

Figura 19. “The Kernel Activity Spaces” o espacios de actividad del núcleo.



Fuente: OMG, 2015, pág. 19.

Respecto a las *habilidades necesarias*, el núcleo establece un conjunto de competencias “clave” requeridas por cada área, las cuales son indispensables para el desarrollo de las actividades de ingeniería de software. Esta parte de ESSENCE, permite orientar a los editores para comprender cuáles son los nuevos roles profesionales requeridos para la construcción de libros digitales, cabe resaltar que la división de perfiles no requiere necesariamente una persona para cada rol, sin embargo, si implica una formación o preparación en los conocimientos que se requieren para el desempeño de cada habilidad, ver Figura 20 .

En el área de *cliente* se requiere un *representante de las partes interesadas* quien debe contar con habilidades para reunir, comunicar y balancear las necesidades del resto de los interesados, pues es quien representa y presenta los puntos de vista de cada uno de estos, al resto del equipo y con frecuencia sirve como interlocutor principal. Esta persona o conjunto de personas requerirá un manejo básico de lenguaje técnico y amplio conocimiento del lenguaje y terminología en el ámbito en el que el software se requiere (OMG, 2015). Este papel desde el ámbito del libro corresponde al editor formado en el manejo de conceptos técnicos y tecnológicos, especialista en la cadena de valor y ciclo de vida del libro y conocedor de las necesidades del lector y el resto de los interesados en la obra.

Figura 20. “The Kernel Competencies” o competencias del núcleo.



Fuente: OMG, 2015, pág. 20.

Para el área de *solución* es necesario contar con personal con habilidades para el análisis, desarrollo y pruebas. El análisis requiere capacidades para identificar necesidades, encontrar oportunidades y transformar, ambas en requisitos bien definidos y consensados. El desarrollo requiere la capacidad de diseñar y codificar software basado en estándares y normas de acuerdo con el producto del que se trate, con el uso de las mejores prácticas y en acuerdo con el equipo. Y las pruebas buscan la verificación de que el software es utilizable y cumple con lo previsto, por lo que es necesario contar con habilidades para explorar a profundidad y encontrar errores.

El *esfuerzo* es la última área y requiere de dos competencias que permitan dirigir con cordialidad al equipo y, además, mantener un seguimiento y control de la carga de trabajo, de aquí que el *liderazgo* es una parte fundamental pues lo ideal es que los equipos se sientan motivados e inspirados a cumplir con los objetivos planteados y para eso se requiere mucho conocimiento tanto del producto como del equipo y compromiso por parte de quién o quiénes dirigirán al equipo. También se requiere personal que se encargue de la *gestión*, lo cual implica dar seguimiento a los planes, proporcionar los recursos, coordinar los esfuerzos y atender las necesidades del equipo de trabajo.

Como puede verse, ESSENCE prácticamente proporciona un conjunto de recomendaciones para el trabajo en equipo, elementos clave a considerar para la construcción de un producto digital y las habilidades que deben desarrollarse por los miembros del equipo, para que éste satisfaga las necesidades de su usuario.

Además, como ya se identificó, estas prácticas o recomendaciones pueden ser útiles si se aplican en la construcción o edición de libros digitales e incluso útiles para la construcción de cualquier otro tipo de publicación digital, sin embargo, para su adecuada aplicación es fundamental comprender el *dominio del problema* o el ámbito en el que se desenvuelve el libro digital conocido como ciclo de vida y que contempla la relación con todos sus usuarios, es decir los autores, editores, distribuidores, intermediarios y el lector que como ya se vio incluye también a las bibliotecas, ya sea porque estas se encargan del resguardo o porque forman parte los medios de difusión a acercamiento con el lector y también, por lo establecido en el modelo OAIS respecto a la preservación pues se requiere “La negociación con quienes producen los objetos de información, para el ingreso apropiado” al archivo de preservación (CCSDS, 2012, pp. 3-1).

Por todo esto ahora es necesario analizar cuáles son los pasos implicados en la edición de libros digitales primero para identificar cómo las practicas esenciales de Ingeniería de Software son útiles o ayudarán a considerar la inserción de metadatos durante el proceso, que como ya se vio son un factor determinante para la preservación y parte de la calidad del software, de ahí que para esto será necesario conocer cuáles son esos pasos que han cambiado respecto al proceso editorial tradicional.

### 2.3. LA EDICIÓN DE LIBROS DIGITALES Y LA SELECCIÓN DE CASOS PARA SU ANÁLISIS

Como ya se dijo, de acuerdo con Bhaskar, la edición experimenta un transición profunda que deriva principalmente del impacto que la transformación digital ha generado en los paradigmas de comunicación, de ahí que ese mecanismo de comunicación tradicional descrito por Darnton, se ve fuertemente afectado, pues ahora “los libros sólo se imprimen en una etapa muy tardía” y “La mayor parte de las labores implicadas ya se efectúa en formatos digitales” (Bhaskar & Rubio Ruiz, 2017, p. 56) es decir que, no sólo el libro se

publica en formato digital, además nace digital incluso si el destino final fuera la versión impresa.

De lo anterior que la comunicación está dirigida por el uso de los medios digitales durante todo el proceso de edición, sin embargo, desde la perspectiva de Bhaskar este no es realmente el problema, pues a pesar de eso, muchas labores se han logrado equiparar a su equivalente en el proceso editorial tradicional adaptando el uso del medio digital.

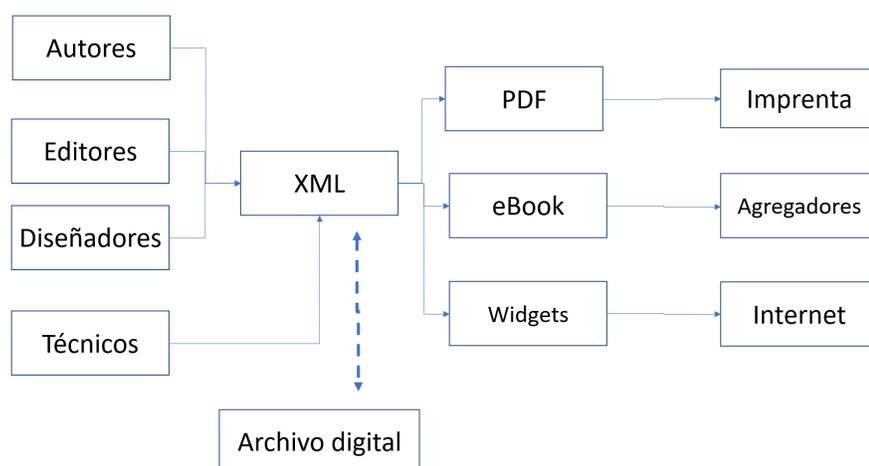
Un elemento que tanto Bhaskar como Gil y Rodríguez mencionan como disruptivo para la edición es Internet y la configuración de redes y comunidades, pues esto crea nuevas formas de interactuar y de cultura, que llevan a la disposición de nuevos parámetros económicos para la producción y distribución de información, generando un fenómeno de convergencia digital que reduce las distancias geográficas, por lo que

Un libro electrónico en la red abierta tiene una capacidad casi infinita de copiarse y compartirse, al instante, en todo el mundo (Bhaskar & Rubio Ruiz, 2017, p. 56)

lo cual, además cambia al contenido mismo, generando, por ejemplo, una desestabilización del derecho de autor y la eliminación de la intermediación, en donde además esto junto con la edición asistida por computadora o autoedición, también conocida en inglés como, desktop publishing, generan incluso dudas respecto a si “cualquiera puede ser editor” o “¿qué es, de entrada, lo que hace que un editor lo sea?” (Bhaskar & Rubio Ruiz, 2017, p. 56). Sin embargo, en este trabajo de investigación no se profundizará en el papel del editor, pero ya se mencionó cuál es la perspectiva que se genera para el proceso de edición con estas dinámicas de autoedición.

Al respecto conviene señalar la forma en que Gil y Rodríguez esquematizan la “digitalización futura” en la edición, que parte de la interacción del autor, editores, diseñadores y personal técnico concentrados en el xml o archivo digital, fuente del pdf, e-book y widgets los cuales van dirigidos a diferentes medios de distribución, la imprenta, internet y los dispositivos como agregadores o contenedores electrónicos, ver Figura 21 (Gil & Rodríguez, 2011, p. 88).

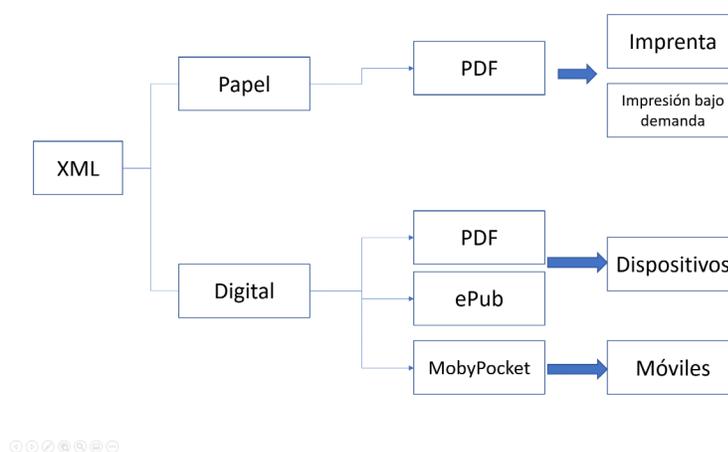
Figura 21. Digitalización futura.



Fuente: Gil & Rodríguez, 2011, p. 88.

Lo anterior, evidencia la disyuntiva que el editor enfrenta respecto a definir el canal de comunicación requerido o determinado para su publicación, esto se ve aún más claro con otro diagrama planteado por Gil y Rodríguez, que refleja lo que también Bhaskar señaló, respecto a que toda publicación nace digital y por lo tanto, el archivo o fichero digital se destina a la versión en papel o la digital, por lo que todo parte de un mismo archivo o fichero origen, sin importar si esta será impresa o digital, ya que el libro puede derivarse en un PDF que se destina concretamente a las versiones impresas, sea por el envío a la imprenta o la impresión bajo demanda o puede construirse como un libro digital en formatos de archivo PDF, EPUB o MobyPocket para dispositivos móviles en donde, además el formato de archivo o fichero puede ser diverso, ver Figura 22 (Gil & Rodríguez, 2011, p. 89) .

Figura 22. Workflow de ficheros.



Fuente: Gil & Rodríguez, 2011, p. 89.

Cabe mencionar, que, respecto a las actividades o acciones envueltas en la edición, Gerardo Kloss, editor y especialista mexicano, ha analizado la complejidad que hay en este proceso y llegado a la conclusión de que lo adecuado es hablar de la edición a partir de las tareas y funciones implicadas, que desde su perspectiva deben contemplar (Kloss Fernández del Castillo, 2020):

[...] las funciones reales [...]: generar ideas, crear contenido, financiar la creación, escribir, reescribir; seleccionar lo que se publica, sugerir cambios, decir que sí, decir que no; coordinar un equipo, financiar la publicación, vender proyectos y hacer marketing previo; corregir y editar textos, hacer cambios sustantivos y accidentales; ilustrar; producir imágenes; manufacturar, diseñar, formar tipografía, capturar, programar, imprimir, encuadernar, controlar calidad; hacer marketing a posteriori; distribuir, diseñar el display; hacer publicidad y ventas, atender puntos de venta; vender en persona, a escuelas, a bibliotecas, por correo, por Internet; definir políticas de descuentos y promociones; dar servicios al lector; evaluar la publicación, ponerla en medios, reseñar, criticar, consagrar, canonizar; recuperar información; archivar; promover el pass along; reutilizar, reciclar, desechar (Kloss Fernández del Castillo, 2020).

De lo anterior, es importante recalcar, que el objetivo de esta investigación no es tratar de definir un proceso editorial para el libro digital, pues el tema requiere un análisis profundo y complejo fuera del alcance de este trabajo, sin embargo, debido a la perspectiva que abordar esta investigación se requiere establecer un proceso de edición que sirva de base o ejemplo para la reflexión que aquí se plantea.

De ahí que, para fines de esta investigación, se presentan casos documentados y publicados de editores, autores o entidades que describen su proceso de edición o recomendaciones para llevar a cabo la publicación digital, los cuales fueron seleccionados durante el desarrollo de este trabajo pues representan cuatro perspectivas para la construcción de libros, las cuales se integrarán en un solo proceso de edición más adelante, con el fin de apoyar en la construcción del modelo objetivo de esta investigación.

Así pues, las perspectivas para llevar a cabo la edición, que a continuación se describen, fueron seleccionadas pues representan los siguientes cuatro supuestos: El primer caso, refleja el proceso empleado por un editor de materiales impresos que contempla la generación de versiones digitales dentro del mismo proceso. Este proceso lo publicó como un manual para editores, la CEPAL<sup>153</sup> un organismo internacional que genera prácticas para apoyar a las comunidades de América Latina.

El segundo caso se trata de la visión o recomendaciones de una autora de textos literarios que presenta elementos orientadores en el camino de la autoedición y además refuerzan el sentido de la esencia de Ingeniería de software, revisada en la sección anterior, de ahí que este caso no tiene como objetivo profundizar en el tema de la autoedición, sin embargo, sirve como un ejemplo más de coincidencia con el trabajo aquí presentado.

El tercer caso es la experiencia descrita como un proceso de edición de un libro digital, seleccionado justamente porque identifica elementos que, desde la perspectiva de los autores consultados, hacen la diferenciación de su proceso con respecto a la edición de un libro impreso.

El último y cuarto caso seleccionado para el análisis, tampoco se refiere a un proceso de edición, sin embargo, expone las recomendaciones publicadas en el sitio web de la DPIG<sup>154</sup>, entidad generadora de estándares a nivel internacional que se ha preocupado por orientar a los editores en la construcción o edición de publicaciones digitales, por lo que este último ejemplo se colocó como referencia de lo que especialistas a nivel internacional en el ámbito de la edición recomiendan y que tienen coincidencia con este trabajo de investigación.

Cabe reiterar que con la selección de los casos que a continuación se presentan y describen, posteriormente se llevará a cabo un análisis para determinar un proceso que sirva a este

---

<sup>153</sup> La Comisión Económica para América Latina y el Caribe (Cepal). Organismo que depende de la Organización de las Naciones Unidas y que es responsable de promover el desarrollo económico y social en América Latina y el Caribe.

<sup>154</sup> Digital Publishing Interest Group

trabajo de investigación como base para el análisis y definición del modelo de integración de metadatos para preservación.

### 2.3.1. GUÍA DE LIBROS ELECTRÓNICOS CEPAL (CASO 1).

Tomaremos como primer ejemplo la guía de libros electrónicos editada por la CEPAL (Ubilla, 2016), que establece medidas destinadas a la producción de libros en formato ePub<sup>155</sup>. Cabe señalar que la CEPAL tiene un proceso editorial que incluye una cadena de actividades para la edición de libros digitales y libros impresos que parte de cuatro actividades en común para ambos tipos de edición, en donde existe: una unidad de control de documentos; continua con una unidad de referencias y terminología; seguida de una unidad de traducción y edición; que concluye en unidad gráfica editorial, ver Figura 23.

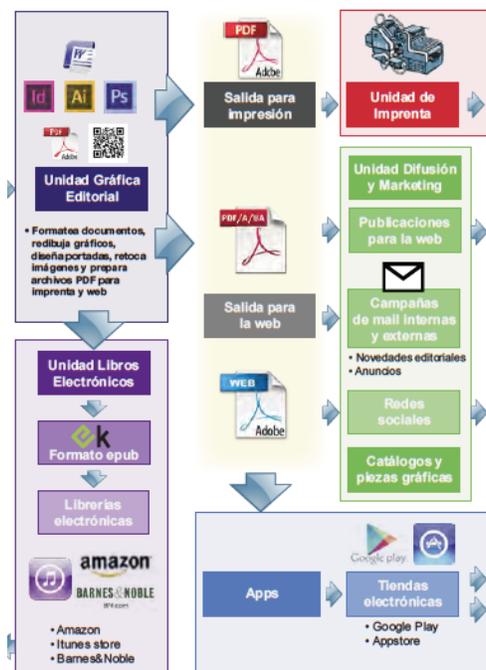
Figura 23. Actividades en común del proceso editorial de CEPAL.



Figura 24. Unidad gráfica editorial de CEPAL.

Fuente: Ubilla, 2016.

Cabe resaltar, que en el proceso de la CEPAL la unidad gráfica editorial construye tres formatos digitales del libro y designa como formato impreso a uno de éstos, que específicamente se produce como archivo PDF, pues es un formato construido para ser distribuido como impreso a su público lector. También genera un archivo en formato PDF/A para distribuirlo por la web y cuenta con una unidad de libros electrónicos sólo para la construcción en formato ePub, ver Figura 24.



Fuente: Ubilla, 2016.

<sup>155</sup> Acrónimo de electronic publication. Se trata de un formato construido con código HTML y adoptado por el Foro Internacional sobre Publicaciones Digitales o International Digital Publishing Forum, que es la entidad encargada de establecer las normas que rigen la industria del libro digital.

CEPAL propone un flujo de trabajo para todos los textos publicados, incluido el formato impreso, pues considera que es muy importante ejecutarlos sin importar cual sea el tipo de distribución del libro, papel o por medios digitales y contemplan cuatro etapas: la primera relacionada con el texto, la siguiente con el diseño, seguido por los estilos y al final algo que llama exportación, ver Figura 25.

Figura 25. Secuencia de elaboración de documentos del proceso editorial de CEPAL.



Fuente: Ubilla, 2016.

- Texto:** El proceso de la CEPAL señala que la edición de los textos debe iniciar con la construcción de estos a través de una herramienta que permita integrar los textos con “estilos jerárquicos”, es decir que recomienda que los autores o creadores desde que inician con la construcción de su obra podrían emplear estos elementos. La definición de estilos contempla establecer características visuales básicas al texto que comprenden el tamaño, la tipografía y el color y en algunos casos, alineación, separación de palabras y uso de diccionarios. Con el uso de estilos es posible generar componentes que le dan estructura al texto y establecer una jerarquía, por esa razón existen estilos para los párrafos, caracteres y objetos. Esta forma de trabajar los textos ayuda a dar formato básico al texto y a clasificarlo pues lo etiqueta para dar continuidad al proceso.

Los estilos de párrafo se aplican a una frase, una oración, una línea o un párrafo desde el comienzo hasta el final de un texto. El estilo de carácter contiene atributos de apariencia para un solo carácter, una sola palabra o serie de palabras que tienen un sentido especial dentro de un párrafo. Es importante mencionar que en esta etapa de la construcción de los textos la estructura que se establece es dirigida a los contenidos a pesar de que se utilicen características visuales de los textos.

De ahí su importancia, pues al definir estilos de acuerdo con el contenido es posible marcar al momento de la construcción las partes conceptuales del texto o del libro, sea cual sea ésta, desde capítulos hasta elementos como títulos de libros, autores, citas, temas o incluso elementos especiales o palabras clave, versos, fórmulas, etc.

Este proceso puede ser de mucha ayuda si se aplica desde que el autor de la obra empieza a conceptualizarla y más aún, si se utilizan herramientas para manejar un etiquetado uniforme y estandarizado, adecuado al tipo de material del que se trate.

- **Diseño:** La siguiente etapa es la construcción visual o formación del libro. Esta normalmente debería llevarse a cabo, una vez que el texto original está estructurado usando los estilos específicos para los párrafos y los caracteres que lo componen. Sin embargo, si esto no sucede así porque el autor o creador de la obra no lo consideró, entonces se sugiere que durante este proceso se realice la estructuración de los textos a partir de la creación de estilos. En este proceso se comienza a trabajar con la formación visual de cada página y elementos que componen el libro: portada, colofón, contraportada, etc. para darle la forma visual que se desea, es por eso, que el proceso implicará la creación de nuevos estilos. Y será posible establecer relaciones entre estilos y hacerlos dependientes unos de otros, además de crear estilos para objetos como imágenes, tablas, etc. El elemento final de esta etapa es la selección adecuada para cada estilo del idioma del diccionario con el cual trabajará cada sección del texto o estilo.

Es importante señalar que en esta etapa se prepara el libro para distribuirse de una forma impresa y CEPAL añade las siguientes dos etapas para la formación del ePub.

- **Estilos:** Durante este proceso se prepara al libro para convertirse en un formato ePub para su distribución en lectores electrónicos. Cabe señalar que este tipo de libro se construye a partir de código XHTML<sup>156</sup>, el cual es un lenguaje de marcado de textos que en su codificación establece una estructura jerarquía y características básicas de apariencia al texto, sin embargo, requiere del uso de archivos CSS<sup>157</sup> u hojas de estilo para el manejo de características más sofisticadas en los textos que incluso pueden implicar la integración de zonas de interacción con el lector. Estas hojas de estilo tienen las mismas características que los estilos que se han definido desde la composición del texto original, por lo que es posible trasladar la mayoría de las características previamente definidas en los estilos a archivos CSS, casi de forma automática, de ahí la importancia de definir estos desde el texto original.

A pesar de que la aplicación de estilos desde las primeras versiones del texto posibilita el traslado de estilos, siempre es necesario realizar una revisión del código para identificar elementos que requieran especificaciones especiales por tratarse de un despliegue en soportes electrónicos, además de la incorporación de las fuentes, pues si los formatos tipográficos no están bien armonizados el traslado de archivos entre las diferentes computadoras con distintos sistemas operativos provoca problemas en la producción. CEPAL recomienda para el intercambio de datos entre

---

<sup>156</sup> Extensible Hypertext Markup Language

<sup>157</sup> Cascading Style Sheets

distintos programas e incluso entre distintos sistemas operativos el uso de tipografías tipo Opentype<sup>158</sup>.

- **Exportación:** De acuerdo con CEPAL, este último paso es el que permite convertir el archivo preparado para impresión a formato ePub o cualquier otro formato, sin embargo, señala que es muy importante que al menos, desde el diseño o construcción visual se siga un orden, pues además de los estilos, existen elementos como cajas de texto u objetos que deben apegarse a una estructura, para que estas no se pierdan durante la exportación. Otro elemento importante aquí es la inserción de metadatos que debe hacerse antes de la exportación. Además, para el caso de los ePub existen ciertos criterios que deben cuidarse de acuerdo con CEPAL, pues son exigidos por tiendas de libros electrónicos como Amazon y Appstore:
  - Los nombres de archivo html no deben contener caracteres especiales como acentos, mayúsculas o espacios en blanco.
  - Las tipografías deben estar descodificadas OEBPST.
  - Todos los enlaces deben ir precedidos por “http://www”
  - Los metadatos deben examinarse con diccionarios.
  - Debe contener un índice estructurado.
  - Las imágenes .jpg deben estar optimizadas.

El proceso de exportación exige que antes de llevarla a cabo se revise las características y restricciones que exige cada formato de exportación para que, al momento del traslado, no se pierdan características del libro.

Como se pudo ver, el proceso de CEPAL hace mucho hincapié en la estructura de los documentos desde su construcción, ya que eso desde su perspectiva, reduce trabajo doble y busca cumplir con requisitos que cada formato de archivo exige, para su buen funcionamiento, además señala la importancia de construir el texto desde que se concibe con una estructura preparada para edición en formato digital.

---

### 2.3.2. EDICIÓN DIGITAL DE TEXTOS LITERARIOS (CASO 2).

El siguiente ejemplo se toma del artículo publicado por González-Blanco que habla de la edición digital de textos literarios, con una perspectiva de autoedición, en donde asegura que las disciplinas humanísticas están rezagadas en el uso de la tecnología con respecto a otras áreas de estudio, por dos razones, primero por la utilidad inmediata, ya que la tecnología, para los autores de este tipo de obras, es sólo de difusión y aprendizaje y no otorga una rentabilidad económica, y la segunda es el perfil de este tipo de creadores, se caracteriza por una escasa interdisciplinaria ya que su método de trabajo es

---

<sup>158</sup> Formato de tipos de letras escalables que permiten su visualización en diversas plataformas o sistemas operativos de dispositivos y equipos de cómputo.

principalmente autónomo y en lo general tiene poco presente en sus métodos de análisis el uso de la tecnología (González-Blanco, 2016).

Por lo anterior, la autora menciona que a pesar de que existen plataformas e información de propuestas para la edición de textos con fines de compartir e interoperar, se vuelve complicado para creadores de textos literarios, conocer y disponer de estos estándares de forma consolidada que les permita entenderlos y usarlos (González-Blanco, 2016).

González-Blanco, menciona a Tapas-TEI, como un ejemplo de comunidad en el ámbito digital, que busca crear un repositorio de código TEI<sup>159</sup> (González-Blanco, 2016), un lenguaje de marcado basado en SGML pensado para el intercambio información textual almacenada en un formato electrónico, específicamente para textos literarios, que hace énfasis en la estructura semántica además de conservar el aspecto visual del texto impreso en un ambiente web, como finales de línea, espacios en blanco, finales de página entre otras características (Burnard, 2001).

Además, González-Blanco menciona plataformas como: TextGrid, espacio virtual alemán para investigación en humanidades, o Ediarum un entorno también alemán para la transcripción de manuscritos con TEI, eLaborate proyecto holandés que proporciona herramientas para edición en TEI, Muruca usada para la creación de biblioteca digitales, sin embargo, asegura que no existen proyectos similares para hispano hablantes, lo cual desde su perspectiva, dificulta para los creadores de obras de esta naturaleza el manejo y construcción de sus publicaciones, de ahí que resalta la necesidad de arduos procesos de capacitación y autoaprendizaje para los autores de obras literarias en el ámbito digital, principalmente hispano hablantes (González-Blanco, 2016).

De lo anterior González-Blanco plantea un conjunto de recomendaciones en cuanto a la autoedición digital de textos literarios(González-Blanco, 2016):

- **Realizar un trabajo “de campo” o de “de navegación”:** para analizar con detalle las diferentes plataformas y herramientas como modelos o ideas inspiradoras de lo que el editor desee hacer.
- **No empezar de cero.** Reconocer el desarrollo de la tecnología aplicada en la edición de textos permite adaptar necesidades sobre un recurso ya existente. Toda propuesta de empezar de cero supone una pérdida de tiempo, aunque también implica un aprendizaje.
- **Conocer y aplicar estándares existes.** Con el fin de fomentar la interoperabilidad entre distintos proyectos y la preservación del contenido digital. En el campo de la edición de textos literarios el modelo TEI es esencial.
- **Establecer una finalidad o alcance.** Se requiere tener muy claro a dónde se quiere llegar, elaborar un prototipo del proyecto que el editor tiene en mente,

---

<sup>159</sup> Text Encoding Initiative. Permite hacer el marcado de textos identificando partes que se refieren al contenido y estructura del texto.

materializándolo y compartiéndolo con los miembros de su equipo. La improvisación y las modificaciones en terrenos digitales son costosas en tiempo, dinero y trabajo, y aunque siempre son necesarias, minimizarlas debe ser el objetivo.

- **Publicar para web.** Realizar ediciones en otros soportes como el CD o aplicaciones de escritorio reduce enormemente las posibilidades de interconexión entre proyectos. La web es el lugar ideal de difusión de contenidos y datos. Busque escribir en un lenguaje como HTML, XML, TEI o XSLT use RDF, metadatos o Linked Data para que sus publicaciones tengan una mayor visibilidad.
- **Usar formatos más accesibles que el PDF.** La utilización de lenguajes basados en XML permite amplias posibilidades de explotación y transformación del texto, ya que establece estructuras semánticas internas dentro de los documentos propiciando el intercambio de datos en la red.
- **Publicar en acceso abierto.** La realidad es que los textos que más se leen son los que se encuentran abiertos en internet. El uso de Creative Commons es una alternativa.
- **Adaptarse a nuevas formas de trabajo.** Los nuevos procesos de edición digital implican nuevas formas de trabajo con dos características importantes: el trabajo en equipos interdisciplinarios, en continuo diálogo entre editores y técnicos, y la colaboración o modelo del crowdsourcing que incorpora varios campos de conocimiento.
- **Reducir la brecha digital.** Ofrecer información y formación sobre posibilidades y perspectivas tecnológicas a los humanistas encargados de la edición al nivel textual.
- **Establecer un proceso claro de edición digital.** La transparencia del proceso de elaboración sirve de guía a los participantes en la edición.
- **Ampliar posibilidades de lectura.** La edición digital debe ir mucho más allá de visualizar sólo el texto, por lo que es recomendable ofrecer más opciones de visualización del contenido, de variantes, de notas aclaratorias de tipo lingüístico o explicativo, e incluso la posibilidad de disponer del manuscrito o impreso original junto al texto a modo de imagen digitalizada.
- **Otorgar al lector la propiedad del texto.** Si bien la publicación conserva intacta su estructura, es posible anotar, convertir, visualizar, adaptar a nuevos formatos, trasladar y luego recuperar según las necesidades del lector.
- **Ampliar las expectativas de los textos.** Los textos editados y etiquetados con lenguajes como TEI presentan nuevas perspectivas y mucho trabajo para humanistas, libreros, filólogos y bibliotecarios. Además de numerosas aplicaciones útiles para instituciones como academias, bibliotecas, archivos y museos.

En conclusión, lo que plantea González-Blanco es un conjunto de recomendaciones generales para creadores de textos literarios en un proceso de autoedición y auto publicación, por lo que, como ya se mencionó, este ejemplo no define un proceso editorial como tal y tampoco necesariamente implica la edición de libros, sin embargo, los

planteamientos que hace resaltar el papel del autor en la edición de una obra y la formación que éste requiere como un factor importante desde la concepción y construcción de sus textos, pues como ya se vio en el ciclo de vida del libro la edición comienza con el autor y en los libros digitales su participación puede implicar conocimientos adicionales o incluso una formación que en la edición tradicional no era necesaria.

### 2.3.3. INTRODUCCIÓN A LA EDICIÓN DIGITAL (CASO 3).

Este ejemplo se tomó del libro descrito por Ordoñez y Galina, respecto al proceso empleado para la edición digital, quienes en primera instancia señalan que, el proceso editorial depende de la arquitectura de la información, la navegación y la interfaz (Ordoñez & Galina Russell, 2007, p. 49) pues con esto, se establece el correcto funcionamiento de la publicación para con su lector.

Además los autores, señalan que para iniciar un proyecto de publicación es necesario contar con un equipo calificado y sugieren la existencia de las siguientes áreas de producción (Ordoñez & Galina Russell, 2007, pp. 65-67):

- **Área de diseño.** Se encarga de definir la interfaz gráfica de la publicación con la cual el usuario va a interactuar, además de establecer el formato de archivo(s) que se utilizarán, por ejemplo, swf<sup>160</sup> o epub.
- **Área editorial.** Área multidisciplinaria puede encargarse de digitalización, corrección de estilo, edición o retoque de imágenes, generación de metadatos y organización de los archivos que componen la publicación. Sin embargo, el trabajo realizado, depende del tipo de tipo de diseño o formato de archivo empleado para la publicación y su funcionalidad.
- **Área de programación.** Encargada de construir la funcionalidad de la publicación entre las cuales están el desarrollo de formularios de búsqueda, versiones de impresión, enlaces a bases de datos y elementos que garanticen la seguridad informática de la publicación.
- **Coordinador del proyecto:** Esta figura, aunque no debe tener conocimientos precisos de cada proceso, deberá conocer las capacidades de las diferentes áreas. Elabora la propuesta de la publicación, desde el nombre hasta la funcionalidad, a quién va dirigida, los aspectos legales y recursos necesarios.
- **Generador de contenido.** Persona o grupo de personas que generan los contenidos que deben cumplir con ciertas normas de calidad de un contenido impreso.

Respecto a las características de la publicación y el formato de publicación, Ordoñez y Galina mencionan que es necesario hacer un análisis, antes de iniciar la edición de los siguientes elementos (Ordoñez & Galina Russell, 2007):

---

<sup>160</sup> Abreviatura de Shockwave Flash, sin embargo, también se conoce como Small Web Format, es un formato de archivo para gráficos vectoriales creado por Adobe Systems.

- El fin u objetivo de la publicación
- El tipo de información que se pretende publicar
- El soporte o medio de publicación de la obra
- El equipo y software disponible para llevar a cabo el proceso editorial
- Los conocimientos técnicos para la integración de las distintas áreas de producción.

Definitivamente, los autores señalan que a partir las cualidades de la publicación se determinar los procesos, lineamientos y costos editoriales necesarios. Pues en las publicaciones digitales existen elementos adicionales que van a acarrear nuevos costos, como son (Ordoñez & Galina Russell, 2007, pp. 67-71):

- Respaldos
- Servidores
- Procesos de digitalización
- Hardware y software
- Optimización de formatos o mejora de la calidad del desempeño
- Cotejo o mejoramiento de OCR
- Uso de lenguajes de marcado para la integración de tipografías especiales o símbolos matemáticos
- Pruebas de interfaz y/o de usuario
- Asignación o generación de metadatos
- Mantenimiento de la publicación posterior a su puesta en distribución

Y concluyen que el proceso de edición en una publicación digital no siempre concluye, pues este puede ser “indefinido” derivado del mantenimiento, que requieren las publicaciones, posterior a su liberación (Ordoñez & Galina Russell, 2007, p. 117).

---

#### 2.3.4. EDICIÓN DE PUBLICACIONES WEB (CASO 4).

En este último caso, se toman en consideración las recomendaciones elaboradas y publicadas por la W3C<sup>161</sup>, para la producción de contenidos en la Web, cabe recalcar que la visión de estas recomendaciones va más allá del libro digital, por lo que son de carácter general y aplicables a cualquier tipo de publicación.

Cabe resaltar que dichas recomendaciones buscan atender el hecho de que “la web se ha convertido en un enjambre de cosas pequeñas que cambia muy rápidamente” (W3C, 2017) por lo cual, resulta necesario establecer un modelo para la creación de este tipo de documento que ayude a reducir su volatilidad.

---

<sup>161</sup> Digital Publishing Interest Group o El World Wide Web Consortium. es una comunidad internacional que trabaja para desarrollar estándares web que permitan desarrollar su máximo potencial. Es dirigido por el inventor de la web Tim Berners-Lee y Jeffrey Jaffe.

Sin embargo, la W3C también considera que a pesar de los evidentes cambios, sigue siendo valioso lograr combinar la edición en papel, una “tradicción antigua y poderosa”(W3C, 2017) con las nuevas expectativas que la plataforma web otorga al libro y en este sentido señala algunas nuevas características importantes para la publicación web:

- Es una colección de uno o más recursos constituyentes.
- Organizados y agrupados con una identificación única.
- Accesible utilizando las tecnologías estándar de la Plataforma Web Abierta.
- Puede considerarse como una manifestación única
- El orden de los recursos constituyentes puede cambiar por la interacción del usuario o las secuencias de comandos.
- Los recursos pueden ser accedidos por el lector directamente a través de la navegación proporcionada por la propia publicación o el agente del usuario
- La publicación se puede empaquetar (y convertirse en un PWP, publicación web empaquetada)
- El empaquetado proporciona una forma conveniente de transmitir y almacenar una publicación, como un epub, por ejemplo.
- Puede ser portátil y estar alojada en algún otro origen.
- Puede estar disponible y funcional mientras el usuario está fuera de línea.
- Tendrá metadatos que describen la publicación en su totalidad.
- Introduce el concepto abstracto de manifiesto, que sirve para llevar información sobre los recursos constituyentes de la publicación.
- Los metadatos y el manifiesto incorporarán información sobre la secuencia y la presentación del contenido.

Sin embargo, la publicación digital está en un punto de inflexión, pues a pesar de que el formato ePub ha sido adoptado para el comercio de libros y este, se basa en tecnologías web, el formato ePub exige características más allá de lo que ofrecen los entornos de creación y más allá de lo que ofrecen los navegadores web para su lectura. Lo cual ha generado que la web en lo general, tenga más funciones que las disponibles para ePub (W3C, 2017).

Por lo anterior, W3C Interest Group plantean los siguientes desafíos para la edición web:

- Los editores deben desarrollar flujos de trabajo, conocimientos editoriales y habilidades tipográficas y de diseño que podrían influir positivamente en la evolución del contenido web en general.
- Una publicación web debe representa una colección de recursos web que, en conjunto, forman una publicación, en lugar de los recursos individuales.
- La publicación debe ser capaz de localizar los recursos consistentemente sin importar el estado de esta, en línea o fuera de línea.

- Una publicación debe identificarse de manera exclusiva y tener uno o incluso varios identificadores y localizadores. Estos pueden ser idénticos, pero también pueden ser diferentes, por ejemplo, un identificador puede referirse a un ISBN específico y otro puede referirse a una copia personal e incluso, el identificador puede ser un método para citar trabajos en línea y permitir citas estables.
- Una publicación web puede constar de muchas partes y es necesario contar con un lugar para almacenar información sobre la publicación y sus componentes. El término manifiesto es el utilizado para esto y debería incluir:
  - Metadatos de publicación, directamente o por referencia.
  - Listado de recursos e información sobre ellos.
  - Secuencia predeterminada de los componentes primarios de publicación.
  - Referencias a tablas de contenido, glosarios y otras ayudas de navegación o pedagógicas.
  - Referencias a manifestaciones alternativas de la publicación.
  - Sugerencias para el agente de usuario, acerca de cómo presentar la publicación.
- Se debe definir una sintaxis para el registro de metadatos en línea básicos<sup>162</sup> que sea independiente del estado de la publicación, que permita el descubrimiento y la recolección de los motores de búsqueda web genéricos, así como los sistemas dedicados de bibliografía y archivo.
- Definir un modelo estándar de API<sup>163</sup> para el acceso a las funciones específicas de las publicaciones a través de programación y así, proporcionar atributos y métodos para acceder a la estructura, el contenido HTML o el contenido textual.
- Mejorar las soluciones para cumplir con todas las expectativas de los editores sobre tipografía, diseño, paginación dinámica y contenido reajutable que soporte explícitamente la adaptación al usuario.
- Transmitir, almacenar, compartir o archivar una publicación web es más fácil si se puede empaquetar como un solo archivo.
- Las publicaciones web deben incorporar un modelo de seguridad que defina las reglas para los estados en línea y portátiles y que se alinee con el empaquetado web.

Como se puede ver, en este último ejemplo tampoco se plantea un proceso de edición para publicaciones digitales, sin embargo, el Digital Publishing Interest Group, trabaja en la creación de criterios que permiten la publicación de cualquier obra en el ámbito digital sobre plataformas web y que, por lo tanto, incluye a los libros digitales. Además, los criterios planteados por este organismo buscan establecer elementos que de una manera eficaz impactan en el funcionamiento de la publicación, pero también en su composición básica y estructura.

---

<sup>162</sup> En la industria de la publicación digital se utilizan ONIX, formato XML o BibTeX.

<sup>163</sup> Application Programming Interface

Además, habla de la importancia de preparar a los nuevos editores, en la integración de metadatos y la definición de un proceso editorial que integre todos estos elementos sin perder las prácticas tradicionales.

#### 2.4. ANÁLISIS DE CASOS PARA LA INTEGRACIÓN DE UN PROCESO DE EDICIÓN.

En este apartado, se describe el proceso de análisis que para esta investigación se definió a partir de la aplicación de una “teoría fundamentada” (Hernández S., 2014, p. 422) en prácticas de ingeniería de software, en las cuales es necesario antes de la creación del modelo, establecer un proceso de edición como base del mismo. De ahí que, fue necesario el “diseño sistemático” (Hernández S., 2014, p. 474) de un proceso editorial, el cual se generó en este trabajo, a partir de los casos descritos anteriormente y basado en la identificación de las siguientes “unidades de análisis” (Hernández S., 2014, p. 474), establecidas como parte de la metodología empleada:

- Prácticas involucradas en cada caso de edición de libro o publicación seleccionado para determinar la organización o procesos determinan tareas sucesivas involucradas en la edición del libro digital.
- Conceptos y/o entidades relacionadas en la cadena de valor del libro digital, que incluye papeles y roles de las unidades de organización que intervienen en la edición de una publicación o libro digital.

Cabe resaltar que además de los casos de edición presentados, en este proceso de análisis también se consideraron elementos señalados por Gil y Rodríguez, Kloss, Darnton y Bhaskar, respecto a los participantes, roles, flujos y funciones que existen en la edición.

De lo anterior y basado en las unidades de análisis antes mencionadas, se tienen las siguientes coincidencias y puntos relevantes:

1. En cuanto a las prácticas involucradas o tareas determinadas en el proceso de edición digital:
  - CEPAL establece cuatro *unidades de trabajo*: la unidad de *control de documentos*, la unidad de *referencias y terminología*, la unidad de *traducción y edición* y la unidad *gráfica editorial*. Las cuatro permiten crear la edición impresa y la edición digital, en donde esta última se construye en formato epub dentro de la unidad *gráfica editorial*, sin embargo, este proceso lo que propone principalmente es la construcción estructurada de los contenidos con el marcado de textos, pues consideran que es la forma más ordenada y con mayores beneficios para generar el archivo ePUB con facilidad. Esto también es señalado por Gil y Rodríguez.
  - González-Blanco, menciona que todo nuevo editor digital debería adaptarse a las nuevas formas de trabajo y establecer un proceso claro, con las siguientes recomendaciones: *Investigación de campo para no iniciar de cero*,

*aplicar estándares, definir un alcance claro y publicar para web.* Esto también lo menciona Gil y Rodríguez.

- Ordoñez y Galina señalan como áreas de producción dentro del proceso: *diseño, edición y programación.* En este proceso se distingue muy bien el trabajo de diseño del de programación, además se plantea que el diseño es diferente dependiendo del formato de la publicación. También existe una actividad de *edición* en donde se incluyen procesos de digitalización además de la generación de metadatos y la organización de los archivos que componen la publicación.
  - La W3C sólo señala entre los desafíos para la edición de publicaciones Web la importancia de que los editores desarrollen flujos de trabajo.
2. En cuanto a conceptos, entidades y roles que intervienen y/o son relevantes en la edición digital:
- CEPAL establece como sustanciales el marcado del texto desde su concepción, la definición de estilos y el trabajo estructurado para el diseño gráfico, pues ayudan a que el proceso de exportación a epub sea transparente y con menos complicaciones, ya que de no ser así todo ese trabajo de uniformidad debe hacerse en el momento de la exportación, lo cual podría ser realmente complicado. Además, integra los metadatos en estas etapas.
  - González-Blanco, menciona que es complicado para los investigadores que editan textos literarios conocer los estándares para edición digital, sin embargo, habla de que el editor digital debería considerar formatos que consideren ampliar las posibilidades de lectura de los textos y priorizar el acceso abierto.
  - Ordoñez y Galina resaltan la importancia de la *arquitectura de la información* y además la presencia de un *coordinador de proyecto de edición digital* quién realiza funciones parecidas a las del editor tradicional. Y, además, señalan la importancia de definir claramente el alcance de la publicación que incluye: el fin u objetivo, tipo de información, soporte en el que se distribuye, recursos con lo que se cuenta: equipo y software, además de las habilidades de las personas que participaran en la edición.
  - La W3C señala algunas características importantes en este tipo de publicaciones: el uso de estándares, metadatos, un identificador único y la existencia de un *manifiesto* o descriptor de cómo se compone la publicación.

De todo esto podemos entonces distinguir e identificar los siguientes elementos relevantes y en común:

1. Es posible considerar que existe una primera etapa en el proceso de edición digital que es muy similar al proceso editorial tradicional, en donde existe:

- a. De acuerdo con la CEPAL, un control de documentos encargado de revisar cuáles son los trabajos editoriales que se tienen recibidos como propuestas, en proceso de aceptación y en proceso de edición propiamente. Además de dar seguimiento a la publicación y en estricto sentido, encargarse de la definición de los alcances de la publicación que, para el caso de la edición digital, tiene las siguientes diferencias:
  - i. En esta etapa del proceso se encuentra lo que mencionan Ordoñez y Galina en cuanto a la importancia de un *coordinador de la edición digital* quién realiza funciones parecidas a las del editor tradicional, quien junto con el Autor define claramente el alcance de la publicación, que incluye: el fin u objetivo, tipo de información y soporte en el que se distribuye, que en conjunto expresan lo que llaman *arquitectura de la información*.
  - ii. Además, la W3C y González-Blanco señalan que en cuanto al uso de estándares además de considerar formatos más accesibles que el pdf, de acceso abierto y que amplíen las posibilidades de lectura y expectativas de los textos. Y González-Blanco lo menciona al decir que se requiere una *investigación de campo para no iniciar de cero*.
  - iii. La parte de estándares expresados por W3C incluye el uso de metadatos, un identificador único y la existencia de un manifiesto que describa cómo se compone la publicación.

En cuanto a la definición de alcances Ordoñez y Galina, mencionan lo importante que es considerar las expectativas de la publicación y el hecho de que depende de los recursos para la edición: que incluye el equipo o infraestructura como computadoras, servicios de telecomunicaciones y software, además del talento humano, que corresponde a las habilidades que se requieren en las personas que participaran en la edición. Recalcando que el alcance debe establecerse desde la primera etapa.

- b. En el proceso de edición tradicional, existe una etapa o momento asignado para la revisión de referencias, terminología y estilos tipográficos, que desde luego también existe en la publicación digital, sin embargo, adicional a esto es necesario considerar lo señalado por CEPAL en cuanto al marcado del texto, la definición de estilos y la inclusión de metadatos, que desde esta etapa puede ir conformándose. Y que Gil y Rodríguez señalan como archivo XML en el flujo de ficheros esquematizado al inicio de la descripción de casos.
- c. Uno de los momentos más exhaustivos o pesados de la edición tradicional podría ser la corrección de estilos y revisión de contenidos que para el caso de la publicación digital también debe darse, pero además de acuerdo con CEPAL el libro digital entra un proceso complejo de marcado de textos,

creación de estilos y definición de metadatos, en donde el uso de estándares será determinante.

- d. Al comparar lo señalado por CEPAL en los incisos b y c, con el proceso de Ordoñez y Galina, es posible concentrar estos dos momentos en la etapa de edición, ya que se habla de la creación de metadatos y captura de información.
2. La penúltima etapa para la publicación digital es la que tiene que ver con la generación del objeto digital, propiamente, que CEPAL lo contiene dentro de una *unidad gráfica editorial* en donde existe una *unidad de libros electrónicos* en donde se obtienen formatos de archivo como epub y que Ordoñez y Galina señalan como *programación* pues se produce la versión del libro digital que contiene funcionalidad como formularios de búsqueda, versiones de impresión, enlaces a bases de datos y elementos que garanticen la seguridad informática de la publicación. También en esta etapa, se construye el manifiesto que menciona W3C y se empaqueta para distribuirse como un objeto.
3. La última etapa es la distribución del objeto y para esto CEPAL señala una *tienda electrónica*, sin embargo, lo importante no es que se comercialice, pues esto puede o no darse, lo importante es definir las formas de publicar o hacer visible el libro ya que esto también forma parte de la definición de alcances de la publicación.

Estos pasos se pueden diagramar para representar las unidades de proceso, la composición de elementos o acciones generales que se desprenden de estas, además de los participantes indispensables, sin que con esto se especifique el número de personas, sólo las habilidades requeridas para llevar a cabo dichas acciones.

Y para eso, es posible utilizar un diagrama de actividades modelado con lenguaje UML, el cual se emplea regularmente en la Ingeniería de Software para representar la ejecución de una secuencia en un procedimiento, el diagrama se compone de cajas rectangulares con los extremos redondeados que contienen la descripción de la actividad, también existen calles o carriles para agrupar actividades en roles o responsabilidades y es posible integrar el flujo de objetos de entrada y salida, representados con rectángulos.

Además, el diagrama de actividades establece una secuencia entre sus componentes o flujo de actividades a partir de flechas continuas que apuntan de forma consecutiva al siguiente componente y usa flechas punteadas para señalar la relación de entradas y salidas de información o resultados de cada actividad, ver Figura 26.

Vale la pena señalar que independientemente del formato o tipo de libro que se trate, las acciones podrían considerarse como generales dentro de un proceso editorial por lo que, para esta investigación se tomará como base para la construcción del modelo que más adelante se presentará.

El proceso define cuatro unidades de trabajo, las cuales comienzan con la creación de una obra intelectual, a partir de la cual se desata el proceso de edición, que se ve compuesto por acciones para cada unidad como a continuación se describe:

- Control y Seguimiento: Alcance y Recursos.
- Edición: que contempla actividades del proceso tradicional (Referencias, terminología, Edición de estilos, Diseño gráfico) y nuevas actividades como la integración de metadatos y el marcado de textos.
- Programación: Funcionalidad, composición, versiones, seguridad.
- Distribución: Medios.

La unidad de *control y seguimiento* es permanente ya que debe acompañar el resto de las actividades del proceso. Además, existen roles o participantes importantes, que como ya se revisó en las secciones anteriores pueden ser muy diversos, sin embargo, en este trabajo de investigación la edición del libro digital contemplará los siguientes roles, de los cuales se hablará más adelante:

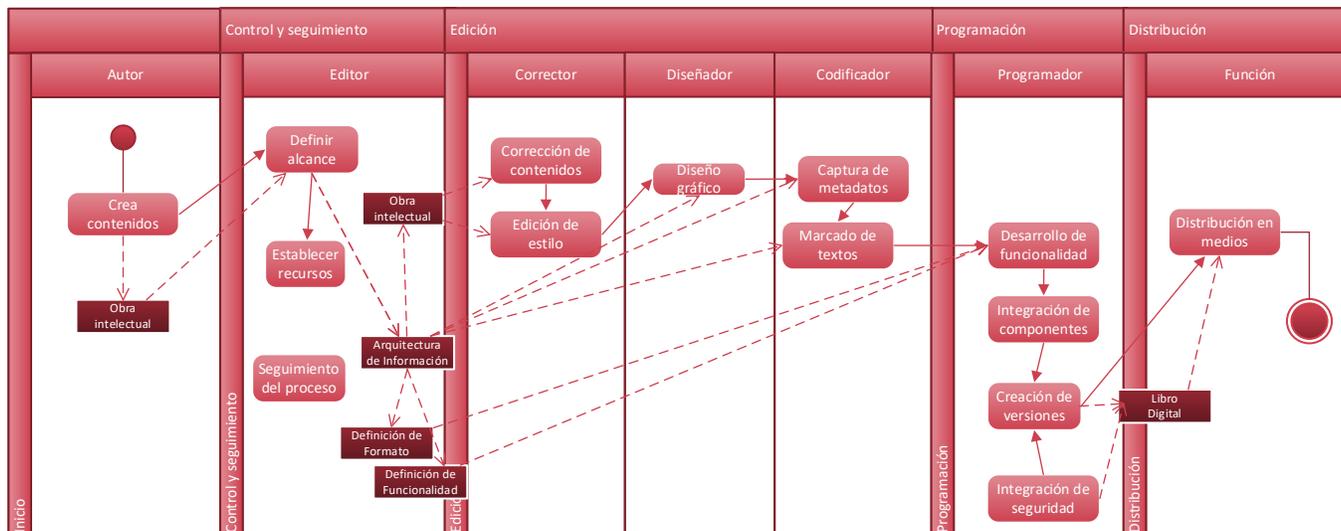
- Autor
- Editor
- Corrector
- Diseñador
- Codificador
- Programador

Cabe resaltar, que el proceso aquí establecido es sólo una propuesta, que integra las recomendaciones e información recabada durante esta investigación pues ayuda a especificar un conjunto de pasos generales que determinan un proceso editorial para libros digitales, sobre los cuales es posible trabajar la siguiente etapa.

Por todo lo anterior, es importante recalcar que este proceso no es aplicado en ningún caso de la vida real o proceso editorial y mucho menos se ha comprobado, por lo que sólo servirá como un supuesto de pasos a seguir y roles básicos para crear publicaciones digitales con el cual se harán comparaciones para la integración de prácticas de Ingeniería de Software

que ayuden a mejorarlo y prepararlo para ser parte del modelo que especificado en esta investigación.

Figura 26. Diagrama de actividades en UML del proceso editorial resultado del análisis.



Fuente: La autora diciembre 2019.

## 2.5. APLICACIÓN DE LA ESENCIA DE INGENIERÍA DE SOFTWARE EN LA EDICIÓN DEL LIBRO DIGITAL.

En esta sección se da continuidad al análisis del proceso de edición del libro, ahora para integrar elementos de ingeniería de software, pues cabe resaltar que la Ingeniería de Software para este trabajo de investigación tiene al menos dos aplicaciones, primero como una herramienta para el análisis y construcción del modelo que se propone, pues permite la descripción de características internas de un sistema (Guasch, 2009) a partir de la aplicación de un lenguaje de representación utilizado para el modelado.

Y la segunda aportación está relacionada con la esencia de la ingeniería de software y sus perspectivas disciplinarias, pues aporta conocimientos en cuanto a prácticas aplicables para el desarrollo de software y por lo tanto, ajustables a los productos digitales o en este caso al libro digital, con lo cual es posible definir “las cosas que siempre se trabajan, las cosas que siempre se hacen y las habilidades que se necesitan” (Jacobson et al., 2012), de ahí que en esta sección, se utilizará para enriquecer el proceso editorial propuesto e integrar nuevos elementos con miras a la preservación a largo plazo.

Para la integración de elementos de Ingeniería de Software, al proceso editorial resultado del análisis anterior se revisarán las cuatro unidades de trabajo definidas:

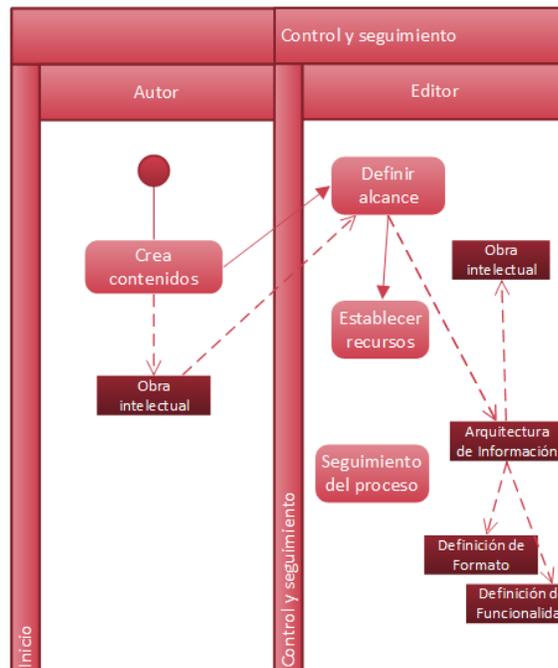
- Control y Seguimiento

- Edición
- Programación
- Distribución

### 2.5.1. ETAPA DE: CONTROL Y SEGUIMIENTO.

La primera unidad de trabajo definida en el proceso editorial de libros digitales es el *control y seguimiento*, que contempla establecer los *alcances* del libro o estructura, funcionalidad o características que tendrá la publicación. Para esto, además es necesario considerar los *recursos* con los que se cuenta o que son necesarios para cumplir con el *alcance*, esto además determina quiénes serán el resto de las personas que participarán en el proceso editorial. El rol más importante aquí es el editor que además requiere de la participación del autor, ver Figura 27.

Figura 27. Control y seguimiento, unidad de trabajo del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

Así pues, esto está muy relacionado con las *alphas* de ESSENCE definidas como *Oportunidad, Requerimientos y Trabajo*, en donde el trabajo se configura basado en los alcances y restricciones definidos en los requerimientos, los cuales se centran en la oportunidad y ésta, es provista por los interesados e involucrados que conciben, consumen y usan el producto digital.

Da ahí la importancia de entender quiénes son los involucrados, pues de acuerdo con ESSENCE es “gente, grupos u organizaciones que afectan o se afectan con el software”

(Jacobson et al., 2013, p. 10). Por lo que, para el caso de las publicaciones digitales, como ya se vio en su ciclo de vida y cadena de valor, los involucrados van desde el autor, editor, corrector, revisor, diseñador y el lector o usuario que incluye a quien lo vende, distribuye, difunde, resguarda, cuida o protege.

Esto también se relaciona con el interés de que una aplicación o software cumpla con lo esperado, conocido como *calidad de software* y descrita en el estándar IEEE 730-2014<sup>164</sup>, como:

grado en que un producto de software cumple con los requisitos establecidos. (SEVOCAB et al., 2021f).

En donde además se recalca que:

la calidad depende del grado en que los requisitos establecidos, representan con precisión las necesidades, deseos y expectativas de los interesados (SEVOCAB et al., 2021f).

Otra definición establecida por el estándar ISO/IEC 25010:2011<sup>165</sup> la menciona como:

el grado en que un producto de software satisface las necesidades declaradas e implícitas cuando se usa en condiciones específicas (SEVOCAB et al., 2021f).

El concepto de calidad, en ambas definiciones está relacionado con el cumplimiento o satisfacción de requisitos o necesidades declaradas o establecidas e implícitas, en donde, además el estándar IEEE 730-2014 recalca que la calidad depende de qué tanto, los requisitos representan las expectativas de los interesados en el software.

También es indispensable considerar que existen al menos cuatro niveles en las necesidades de los interesados (Laporte & April, 2018):

- **Necesidades reales:** las que están en su mente.
- **Necesidades expresadas:** Las que se manifiestan de forma verbal o escrita.
- **Necesidades especificadas:** Las que quedan plasmadas en la documentación o especificación como un requerimiento.
- **Necesidades cumplidas:** Las que después de estar plasmadas en la documentación, fueron verificadas o probadas que se cumplen en el producto de software.

Además, existen factores generadores de errores durante la definición de necesidades y que por lo tanto van a repercutir en la calidad, como (Laporte & April, 2018):

- 1) **La comunicación no efectiva entre interesados y desarrolladores:** Por lo que es necesario uniformar terminología y usar un lenguaje lo menos técnico posible, para lograr que la comunicación se acerque a la realidad de los interesados y evitar malentendidos involuntarios. Además, se recomienda: tomar notas en cada reunión

---

<sup>164</sup> IEEE Standard for Software Quality Assurance Processes

<sup>165</sup> Systems and software engineering--Systems and software Quality Requirements and Evaluation (SQuaRE)

y distribuirlas a todo el equipo; revisar los documentos producidos; ser consistente en el uso de los términos y desarrollar un glosario de términos para compartir con todos los interesados; informar a los interesados del costo de cambiar las especificaciones; utilizar un enfoque de desarrollo que permita aceptar cambios en el camino; numerar cada requisito y llevar un proceso de gestión de cambios para cada uno.

- 2) **Desviación en las especificaciones:** Esto generalmente sucede como una derivación del problema anterior, pues al interpretar incorrectamente un requisito se crea un software según el entendimiento propio, haciendo ajustes sin consultar con los interesados las adecuaciones al nuevo producto, lo cual puede ser originado también por presiones presupuestales o de tiempo. Otro factor importante que desencadena este problema es la falta de supervisión o control durante el proceso de desarrollo, lo cual provoca que los errores se detecten hasta que el producto se usa.

Con base en esto podemos decir que, los principales problemas en la especificación de alcances están en la falta de entendimiento y formas de especificar las necesidades que expresan los involucrados; y en el control y supervisión, para confirmar que las necesidades se están cumpliendo durante todo el proceso de creación del libro digital.

En cuanto a *el control y la supervisión*, estas actividades son parte del esfuerzo para el desarrollo de un software señaladas dentro de las *cosas por hacer* en ESSENCE que de forma general contemplan (Jacobson et al., 2013, p. 22):

- **Preparase para hacer el trabajo:** Alistar el equipo y su entorno de trabajo. Esto implica comprender el trabajo para contratar al personal adecuado.
- **Coordinar actividades:** Coordinar y dirigir el trabajo del equipo. Esto incluye toda la planeación y su adecuación durante el curso del trabajo.
- **Apoyar al equipo:** Ayudar a los miembros del equipo a que se ayuden a sí mismos, colaborar y mejorar su forma de trabajo
- **Rastrear el progreso:** Medir y evaluar el progreso que hace el equipo.
- **Detener el trabajo:** Cerrar el trabajo y transferir las responsabilidades del equipo.

Lo cual significa que el editor o el responsable de la edición digital deberían como parte del *control* comprender claramente lo que se debe de hacer para contratar al personal calificado, además de planear las actividades que cada persona desempeñará, considerando también que fungirá como un apoyo para el equipo pues se encarga de resolver los problemas, evaluar y medir el progreso de la publicación, además de establecer el momento en que el trabajo de la edición termina para así asignar al equipo nuevas publicaciones.

En este sentido, ESSENCE señala dos habilidades importantes, el *liderazgo* y la *gestión*, que son indispensables para *el esfuerzo* que lleva a cabo un equipo de trabajo y que en este caso

lo enfocaremos en el proceso de edición del libro digital. En donde, además, basados en los 5 niveles de competencia que señala ESSENCE, es posible utilizar el nivel básico y adecuarlo para definir puntos de autoevaluación que la persona, que asume este rol, puede ocupar para reflexionar:

- Entiende cuál es el papel profesional que debe desempeñar como líder y gestor de una edición digital.
- Es capaz de responder correctamente a cuestionamientos básicos que corresponden al liderazgo y gestión del proceso de edición de una publicación o libro digital.
- Es capaz de ejecutar las funciones básicas de liderazgo y gestión en la edición de una publicación digital.
- Puede seguir instrucciones y completar tareas básicas de liderazgo y gestión.

En conclusión, es conveniente que el *control y supervisión* del proceso editorial de publicaciones digitales contemple las siguientes recomendaciones de la IS:

**1. En cuanto al control para la definición del alcance y recursos de la publicación.**

- a. Considerar los recursos que se tienen para elaborar la publicación es determinante para definir la complejidad de la publicación por lo que será importante meditar si se cuenta con los recursos humanos, económicos y tecnológicos para hacer realidad las necesidades especificadas. Esto incluye un apoyo y soporte para el editor tradicional, pues debe establecerse si requiere de un líder de proyecto o personal dedicado al *control y seguimiento* de la publicación digital.
- b. Dentro del alcance se encuentra definir la estructura de la publicación y su funcionalidad, en donde es indispensable la participación del editor, autor y un especialista encargado de la publicación digital.
- c. Al definir el alcance de la publicación es posible determinar los tiempos y recursos para las siguientes unidades de trabajo: producción, programación y distribución.
- d. Utilizar un glosario de términos que permita uniformar el lenguaje técnico tecnológico relacionado con el desarrollo del libro digital, para llegar a acuerdos en cuanto a los alcances de la publicación, con lo cual todos los involucrados podrían entender a qué se refiere cada característica funcional y no funcional de la publicación. Esto incluye terminología para la producción, programación y distribución.
- e. Incluir como parte de los requerimientos necesidades generales del público al que va dirigido la publicación, que incluye desde los distribuidores y lectores hasta las entidades encargadas del resguardo de las publicaciones digitales. Para esto es útil definir criterios básicos de calidad en una publicación digital desde la perspectiva del software, que en esencia deberían contemplar el uso de

estándares y prácticas generadas previamente en la organización o por la comunidad de editores, lectores, codificadores, programadores, distribuidores, bibliotecas y/o entidades encargadas del resguardo de publicaciones.

- f. Definir por escrito las características que precisan el alcance de la publicación, para tratar de confirmar que las necesidades reales son igual a las expresadas y especificadas, de tal forma que al final, estas mismas sean necesidades cumplidas. Esta información una vez trabajada, permite describir las características que tiene la publicación y es la primera entrada de metadatos en la elaboración del libro digital.

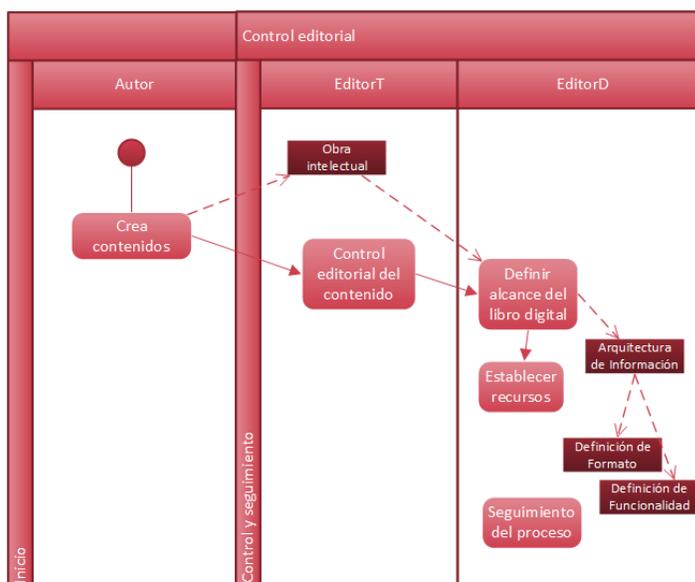
## **2. En cuanto al seguimiento:**

- a. Es una actividad que inicia al momento de establecer el alcance de la publicación y allegarse de recursos para lograrlo, que incluye conformar un equipo editorial con la preparación y/o contratación del personal adecuado que va desde el editor o especialista que dará seguimiento a la publicación digital y los integrantes del equipo como revisores, correctores, diseñadores, programadores e incluso distribuidores de la publicación.
- b. El editor o líder de la edición del libro digital es el responsable directo y el apoyo para el equipo. Pues es el encargado de vigilar el avance y resolver los problemas, buscando la mejor comunicación entre todos los involucrados, resolviendo cualquier malentendido involuntario. Es quien convoca a reuniones para tomar acuerdos y lleva un registro de estos, distribuyéndolos a todo el equipo, elabora el plan y revisa la producción de cada miembro del equipo. También está al tanto y lleva control de los costos de la publicación y los costos generados por cambios inesperados.
- c. El seguimiento es un trabajo colaborativo de todos los involucrados. Pues lo recomendable es que todos los participantes en la edición tengan claro cuál es el alcance de la publicación y el camino a seguir en general, además de sus responsabilidades y recursos en particular, de tal forma que puedan autogestionar el trabajo asignado y así contribuir a una vigilancia de todo el proyecto. Es todo el equipo el que idealmente debiera evaluar desde su experiencia nuevos cambios en la publicación y su pertinencia, de ahí la importancia de que el editor del libro digital mantenga una buena y clara comunicación con su equipo.
- d. La verificación del proceso de producción y de la publicación, incluyen evaluar y medir el progreso del proceso, documentando cómo se hace las cosas para aprender de las buenas y malas prácticas e ir mejorando. Pero además es indispensable comprobar la calidad de la publicación digital como software y para esto es necesario utilizar estándares de calidad desde la definición de alcances de la publicación y verificar su cumplimiento.
- e. El seguimiento debe establecer el momento en que el trabajo de la edición termina, para así dar por concluida la publicación y distribuirla de acuerdo con

los alcances especificados para asignar al equipo nuevas actividades o publicaciones.

De todo lo anterior, se pueden concluir dos cosas: primero que el alcance de la publicación deben considerar desde el inicio elementos que satisfagan necesidades de preservación a largo plazo y segundo, la necesidad de integrar un nuevo rol en la propuesta planteada, que acompañará al editor tradicional para llevar a cabo labores de control y seguimiento además de adquirir nuevas habilidades que le permitan ser capaz de establecer con su equipo, el alcance de la publicación e identificar los recursos que se requieren, al cual para fines de esta investigación se nombrará como *editor digital*, a partir de la propuesta de Gil y Rodríguez, pues esta persona debe conocer el lenguaje del editor y aprender un nuevo lenguaje digital que implica entender la estructura de la publicación o características funcionales e incluso, la importancia que tiene el manejo de estándares para el mercado del texto y los metadatos.

Figura 28. Cambios en unidad de trabajo control editorial del proceso propuesto



Fuente: La autora diciembre 2019.

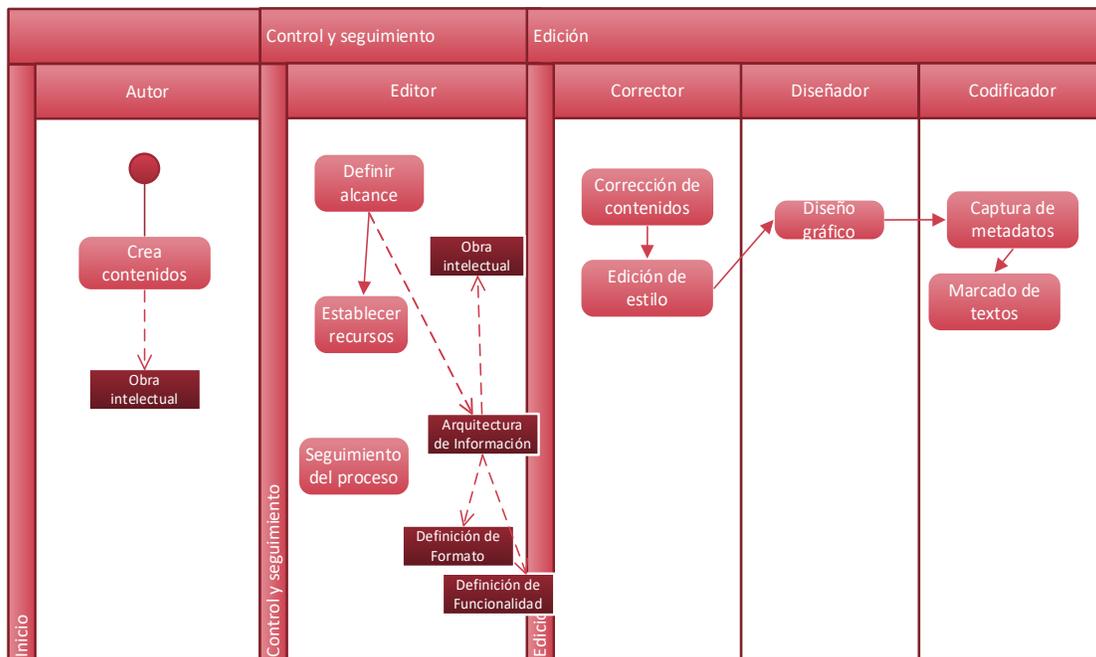
Así pues, la unidad de trabajo *control y seguimiento*, se modifica para ahora llamarse *control editorial*, en donde se establece la participación de dos editores responsables del proceso, *EditorT* o editor tradicional y *EditorD* o editor digital, que en su momento deberán coordinarse para establecer el alcance de la publicación y dar seguimiento al desarrollo desde ambas perspectivas, la tradicional y la digital, ver Figura 28. Además, el seguimiento formará parte de las actividades del *EditorD*, como se verá más adelante.

### 2.5.2. ETAPA DE: EDICIÓN

La *edición* establecida como unidad de trabajo del proceso editorial de libros digitales contempla acciones tradicionales en la edición de libros, de ahí que, los roles de corrector/revisor y diseñador participen como actores importantes.

Además, se tiene considerado un rol identificado como *codificador* pues como se vio en los ejemplos analizados en el capítulo anterior, una parte importante es establecer una construcción uniforme, ordenada y estructurada de los textos con el marcado adecuado y la integración de metadatos en la publicación, sin importar de qué tipo de publicación se trate y si esta será impresa o puramente digital, ver Figura 29.

Figura 29. Edición, unidad de trabajo del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

Sin embargo, cabe aclarar que la cantidad de acciones que integran la *edición* dependen directamente del alcance, funcionalidad y tipo de publicación que se pretende obtener, pues como ya lo señalaban Ordoñez y Galina en uno de los ejemplos revisados en el capítulo anterior, aquí es posible contemplar digitalización, edición de audio y video, ilustración, animación y captura de información en bases de datos o lo que se requiera, no obstante, lo que nunca debe faltar son las partes del proceso tradicional y la codificación para el marcado estandarizado de textos e integración de metadatos.

De ahí, que para fines de esta investigación esta etapa se renombrará como *producción* y será importante recalcar que antes de iniciarla, debe establecerse claramente el alcance de

la publicación desde la etapa anterior e incluir los estándares para el etiquetado del texto y metadatos que se aplicarán.

Así que, para esta unidad de trabajo lo recomendado por ESSENCE está señalado en el *alpha: sistema de software*, pues se refieren la composición de software, hardware y datos que coordinadamente generan un valor para las partes interesadas. En donde las recomendaciones de verificación están orientadas a los siguientes elementos:

- Definir o diseñar la arquitectura del producto.
- Establecer elementos que permitan comprobar que la arquitectura seleccionada se cumple.
- Establecer elementos que indiquen que el producto es usable.
- Definir elementos para saber cuándo el producto está listo.

Es decir que, en esta etapa, además de llevar a cabo la revisión de los textos con los procesos tradicionales, es necesario establecer con base en los alcances de la publicación la *arquitectura de la información o arquitectura del libro* que Ordoñez y Galina mencionaban, pues con base en eso se establecen las estructuras de marcado y la composición o construcción de la publicación que incluye el software, hardware y datos necesarios, además de la realización, programación o codificación de las funcionalidades y los elementos que permitirán realizar la comprobación de que cumple con el objetivo esperado (OMG, 2015).

De ahí que ESSENCE señala que, para la capa de *solución*, que en este caso se refiere al libro digital que se pretende construir, es necesario contar con habilidades de *análisis, desarrollo y pruebas*, en donde las habilidades de *análisis* corresponden a los conocimientos adquiridos para entender cuáles son las oportunidades que una publicación digital otorga o debería otorgar a sus autores, lectores, distribuidores y entidades encargadas de su resguardo y preservación e integrar las características que permitan cumplir con dichas expectativas.

Las habilidades de desarrollo se refieren a las capacidades para programar las funcionalidades que fueron especificadas como parte de la arquitectura de la información o arquitectura del libro. Y las habilidades de pruebas, se refieren a las capacidades para verificar la usabilidad, estructura y funcionalidad de la publicación.

De ahí que, para esta etapa, independientemente del alcance de la publicación, es necesario considerar las siguientes acciones derivadas de la IS:

**1. En cuanto a la arquitectura de la publicación:**

- a. Establecer por escrito la definición de la arquitectura, componentes y funcionalidades.
- b. Establecer la *arquitectura de la publicación* basada en el tipo, funcionalidades y cualidades que esta debe cumplir para el autor, lector, distribuidor y entidad de

resguardo. De ahí que la arquitectura puede incluir desde sólo la estructura de marcado del texto hasta la estructura de bases de datos cuando así sea necesario. Siempre considerando que la definición de la arquitectura de la información es parte importante para la preservación a largo plazo, pues como ya se vio es un elemento señalado constantemente como parte del libro y que forma parte esencial del mismo.

- c. Establecer o definir por escrito las herramientas de software y hardware que se requieren para el marcado y codificación de la publicación, pues esto implica hacer el análisis necesario para determinar el uso o de elementos que ayuden a la preservación de publicación digital a largo plazo.
- d. Para esta labor es necesario contar con personas calificadas para llevar a cabo el diseño de la arquitectura de la publicación, lo cual contempla determinar las actividades adicionales que deben integrarse a la producción como digitalización, edición de audio, video, animación, ilustración e incluso el diseño de bases de datos.

**2. En cuanto al marcado del texto y metadatos o codificación:**

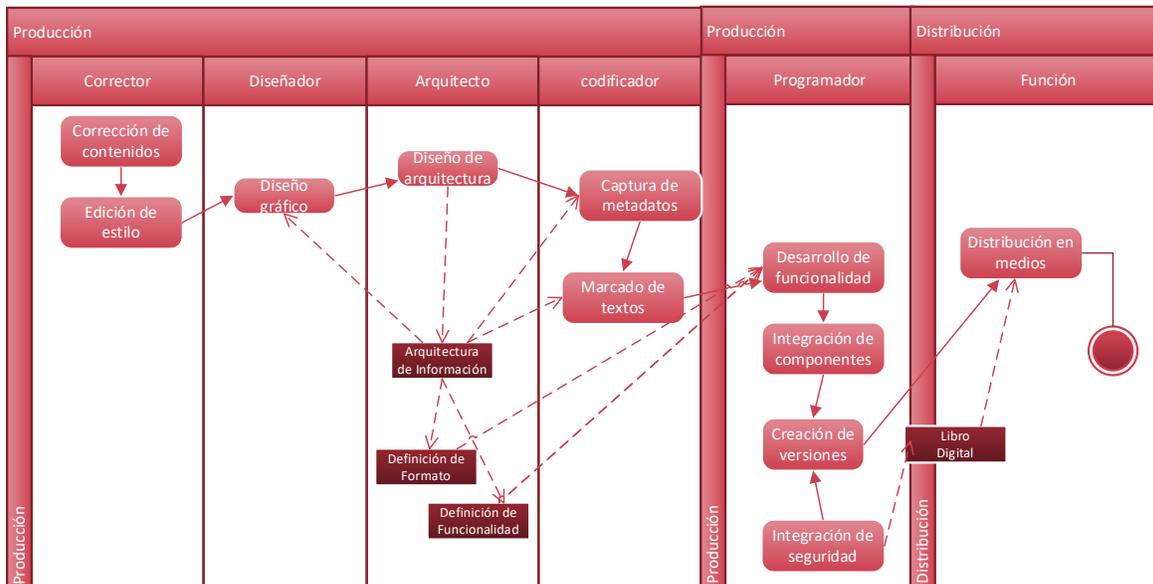
- a. Establecer una estructura de metadatos de acuerdo con el tipo de publicación que se trate y la arquitectura definida, la cual al menos deberá cubrir dos aspectos importantes: la descripción de la obra y la descripción de la estructura de la información que contempla la estructura de archivos y carpetas, pues esto es parte importante para preservación a largo plazo.
- b. El marcado de textos acompaña a la inserción de metadatos y es labor del codificador o codificadores, sin embargo, es trabajo del arquitecto de información crear la estructura adecuada al tipo de publicación.
- c. La documentación de la arquitectura también es labor del codificador y debe estar contenida en una estructura de metadatos adecuada. Sin embargo, cuando se trate de bases de datos, es posible que la inserción y definición de metadatos de arquitectura estén del lado de los programadores, por lo que el arquitecto de la publicación define el alcance de las responsabilidades de estos con base en el diseño de la arquitectura.
- d. Se recomienda, usar esquemas de metadatos estandarizados de acuerdo con el tipo de información que se quiera almacenar, pues como se vio en el primer capítulo de esta investigación, existen diversos esquemas y cada uno obedece a cierto tipo de información, por mencionar algunos, Dublin Core, BiBframe, MODS, METS, Marc21 y ONIX.

Como puede verse, la *producción* en el proceso editorial propuesto sólo considera una parte de las actividades que intervienen en la construcción de la *solución* de acuerdo con ESSENCE, sin embargo, las actividades faltantes se ubican en la unidad de trabajo *programación*, que se verá más adelante, por lo que está se puede adecuar aún más.

Por lo pronto, vale la pena concluir que todo lo anterior sugiere modificaciones en el diagrama de la unidad de trabajo *producción*, pues es necesario incluir un rol de arquitecto, quien como ya se dijo, se encarga de diseñar la estructura lógica y física de la publicación con base en el tipo de publicación y las funcionalidades que esta cumple y que fueron definidas en el alcance, ver Figura 30.

El arquitecto también es el encargado de diseñar el modelo de datos para la integración de metadatos, que como ya se mencionó debe contener tanto la descripción de la obra como la descripción de la arquitectura y basarse en el manuscrito revisado y corregido en los procesos tradicionales. De ahí que este también defina el alcance de las labores de programación, codificación y diseño.

Figura 30. Cambios en la unidad de trabajo Producción del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

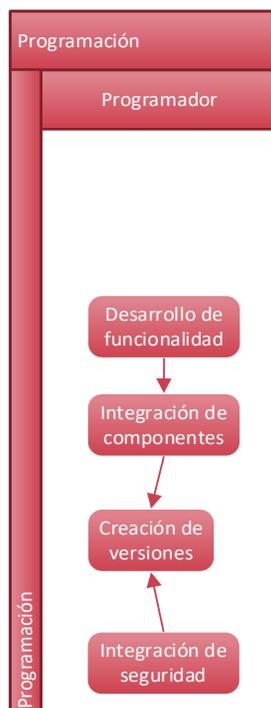
El arquitecto también define las herramientas de trabajo empleadas para el codificador y en algunos casos para el programador, ya que es importante determinar cómo será el ingreso de metadatos para describir el contenido de la obra y la arquitectura, de ahí que es posible que algunos metadatos estén a cargo de los programadores en ciertos momentos.

### 2.5.3. ETAPA DE: PROGRAMACIÓN

La *programación* es una unidad de trabajo que va de la mano con la *producción*, sin embargo, cuando se trata de publicaciones con poca o nula funcionalidad podría pensarse que esta etapa no es necesaria, lo cual no es cierto, pues también incluye la preparación del libro para su liberación y entrega, por lo que es una actividad indispensable en todo proceso editorial de publicaciones digitales para verificar la calidad del software o técnica que

incluye características de funcionalidad, compleja o tan sencilla como la navegación entre enlaces o lingas dentro y fuera de la publicación, además que implica revisar las versiones para confirmar el formato de archivo y la composición del entregable o libro, respecto a la seguridad y metadatos por ejemplo, ver Figura 31.

Figura 31. Programación. Unidad de trabajo del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

En este sentido, vale retomar lo señalado con respecto a ESSENCE en el *alpha: sistema de software*, en donde, además de diseñar la arquitectura se recomienda:

- Establecer elementos que permitan comprobar que la arquitectura seleccionada se cumple.
- Establecer elementos que indiquen que el producto es usable.
- Definir elementos para saber cuándo el producto está listo.

Para lo cual, antes de profundizar en esta etapa del proceso editorial, es conveniente retomar el concepto de calidad de software y revisar las normas ISO/IEC 9126<sup>166</sup>, ISO/IEC 14598<sup>167</sup>, e ISO 25000:2005 SQuaRE<sup>168</sup>, de las cuales se desprenden los siguientes ocho *atributos de calidad de software* (Boehm et al., 1976):

<sup>166</sup> Calidad del producto de software

<sup>167</sup> Evaluación del producto

<sup>168</sup> System Quality Requirements and Evaluation

- **Funcionalidad:** Características que describen las funciones, tareas u objetivos que cumple el producto de software.
- **Fiabilidad:** Capacidad para recuperarse de fallas y mantenerse en servicio.
- **Usabilidad:** Características que facilitan el uso, con un mínimo de esfuerzo.
- **Eficiencia:** Capacidad para responder con la velocidad apropiada, con el uso recursos adecuados.
- **Mantenimiento:** Características que permiten mantenerlo o mejorarlo con el mínimo esfuerzo.
- **Portabilidad:** Capacidad para adaptarse a diferentes entornos.
- **Compatibilidad:** Capacidad para intercambiar información con otro software que incluye: la coexistencia y la interoperabilidad.
- **Seguridad:** Grado en el que se protege, que incluye: confidencialidad e Integridad de datos, responsabilidad de los usuarios, pertenencia y autenticidad de la información.

Todos estos atributos podrían cumplir con necesidades declaradas por los interesados o ser establecidas por estándares o políticas aplicadas de acuerdo con el tipo de producto digital del que se trate, pues están relacionadas con su funcionamiento, con su estructura y con las características que cumple para interactuar con otras aplicaciones y por lo tanto pueden ser evaluadas y verificadas durante la construcción del producto digital y en este caso durante el proceso editorial.

De ahí que, es posible afirmar que todo software independientemente del propósito que cumple, puede ser evaluado por sus atributos o características básicas y estas también pueden verse impactadas por factores generadores de errores que existen durante el diseño de la arquitectura, la codificación, las pruebas y la documentación por lo que es necesario tomar en consideración lo siguiente (Laporte and April, 2018):

- 1) **Problemas en la arquitectura y diseño.** Estos suceden durante la traducción de requisitos en especificaciones técnicas. Están derivados de la falta de comunicación o de la carencia de un rol específico para esta actividad, los síntomas de estos problemas son: la falta de especificación de datos primarios y de funciones de procesamiento de datos esenciales, la falta de especificación para el procesamiento de errores, excepciones y operaciones ilegales.
- 2) **Problemas en la codificación** (incluido el código de prueba). La mayoría de estos ocurren durante la construcción y van desde una elección inadecuada del lenguaje de programación o herramientas hasta la falta de especificación de convenciones para la codificación o incluso falta de pericia en construcción de algoritmos, falta o mala validación de datos, tratamiento inadecuado de tipos de datos, manejo incorrecto de índices, manejo ineficiente de excepciones además de una mala documentación del código.

- 3) **Problemas por el incumplimiento de los procesos / procedimientos.** Estos están ocasionados por una falta o inadecuada aplicación de procesos, que finalmente repercute o va generando cualquiera de los problemas antes mencionados. De ahí que es necesario establece una metodología ya sea interna o basada en estándares que describa claramente pasos a seguir, entregables, formatos, estándares a utilizar para la documentación, diseño, codificación y pruebas. Los síntomas derivados de la existencia de este tipo de problemas se evidencian por la falta o inadecuada documentación que provoca dificultad para entender, usar, probar, mantener e inclusive remplazar el producto.
- 4) **Problemas por revisiones y pruebas inadecuadas.** El proceso de prueba tiene por objeto identificar fallas para corregir defectos, por lo que esta actividad debe ejecutarse con el suficiente rigor para cumplir con su cometido. Los puntos críticos están en el alcance de los planes de prueba que además deben contemplar la revisión de la documentación o entregables intermedios en contraste con las pruebas del código y asegurar que contemple todas las funcionalidades. Además, es necesario dar seguimiento a los defectos para llevar un registro completo y así certificar que se corrijan. El factor que más afecta al proceso de pruebas es la falta de tiempo, ya sea por el retraso en las fases anteriores o por una planeación que no contemple esta actividad. Los síntomas que hablan de este tipo de problemas son: funcionalidades sin probar, defectos no corregidos, tiempo insuficiente para pruebas y la falta de comunicación de los defectos encontrados.
- 5) **Problemas con la documentación.** Es un problema común que la documentación del software sea obsoleta o incompleta y también es común, que los equipos de desarrollo no quieran invertir tiempo en su revisión o actualización. Por lo que la obsolescencia de la documentación se da por cambios durante el ciclo de vida del software, que no se reflejan en la documentación.

Con todo esto y como se vio en la etapa anterior, las acciones contenidas en esta unidad de trabajo se relacionan mucho con el diseño de la *arquitectura de la publicación* pero no sólo con las labores de programación como lo definen Ordoñez y Galina, además existe un proceso de revisión y verificación, por lo que tal vez conviene asignar un nuevo nombre y reestructurar las actividades de esta unidad de trabajo, de tal forma que incluso también modifique la etapa de *producción*, acercando más lo que es puramente la programación con el proceso de producción y estableciendo una unidad de trabajo centrada en el *control de calidad del software* o técnica-tecnológica de la publicación, que además necesita de un responsable para dichas acciones y que desde luego se basa tanto en los alcances como en el diseño de la arquitectura de la publicación, establecidos al inicio del proceso editorial.

De ahí que, para esta etapa, es necesario considerar las siguientes acciones derivadas de la Ingeniería de Software:

1. **En cuanto a calidad de la publicación o software:**

- a. La documentación elaborada para la definición de la arquitectura en la etapa anterior es el principal instrumento de trabajo aquí, pues permitirá comprobar que la arquitectura seleccionada para la publicación se cumple, de ahí la importancia de encontrar la manera de integrar esta información como metadatos en la publicación, que permita una actualización constante.
- b. Con base en los atributos de calidad es necesario que la evaluación técnica-tecnológica de la publicación considere los siguientes elementos que se tomaran de los criterios específicos en la documentación elaborada para el diseño de la arquitectura de la publicación: Funcionalidad, Fiabilidad, Usabilidad, Eficiencia, Mantenimiento, Portabilidad, Compatibilidad y Seguridad.
- c. La funcionalidad se refiere a la interacción que la publicación contempla con sus usuarios o lectores. La verificación puede hacerse simulando el uso y comparando los resultados de acuerdo con lo especificado en los alcances y arquitectura de la publicación. Pero también debería ser considerada en la funcionalidad, la contención de metadatos que auto documenten la publicación y permitan su interacción en ambientes de preservación digital.
- d. La usabilidad es un atributo sumamente especial, ya que involucra características en la publicación relacionadas con la facilidad de uso y lectura, por lo que se relaciona con las estrategias aplicadas para integrar la funcionalidad y las decisiones tomadas en cuanto al resto de los atributos.
- e. La fiabilidad, eficiencia, mantenimiento, portabilidad, compatibilidad y parte de la seguridad son atributos que deben contemplarse desde el diseño de la arquitectura de la publicación, de tal forma que las herramientas que se determinen tanto para la construcción como para el uso de la publicación deben apegarse a estándares y ser evaluados en su vida útil, para evitar la creación de publicaciones con alto grado de obsolescencia desde su nacimiento.
- f. La fiabilidad, compatibilidad, portabilidad y mantenimiento en el caso de las publicaciones digitales son atributos estrechamente relacionados con los metadatos y el marcado del texto, por lo que nuevamente es necesario recalcar que el uso de estándares de metadatos adecuados al tipo de publicación contribuye en la preservación de esta.
- g. La seguridad en el caso de las publicaciones digitales contempla las estrategias empleadas durante la construcción y aquellas implementadas para el uso, pues por un lado implica asegurar la integridad del material que va desde contener los metadatos descriptivos y técnicos necesarios, hasta contar con las restricciones necesarias para cuidar que se cumpla con los derechos otorgados a los lectores o usuarios de la obra.

**2. En cuanto a la calidad del proceso editorial:**

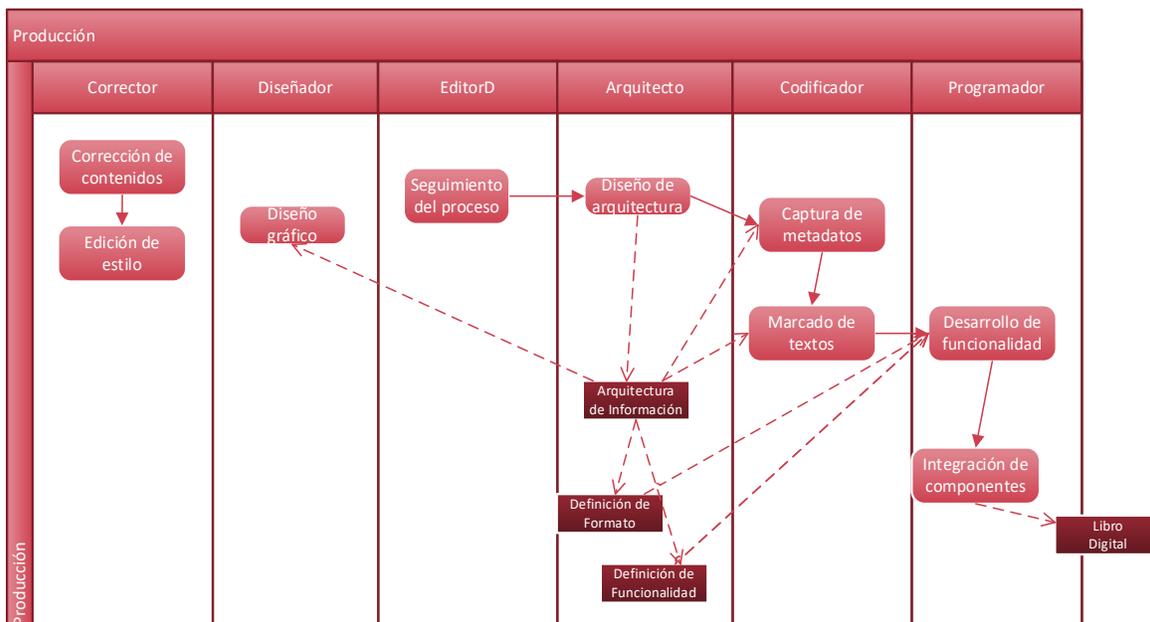
- a. Parte importante del proceso es generar la documentación necesaria que alimenta las necesidades de información durante cada etapa, por lo que a pesar de que se menciona hasta este momento, es necesario que el proceso considere

desde el inicio, un medio estandarizado de registro tanto de los alcances esperados como de la arquitectura definida, pues esto permitiría que esta información sea parte de la publicación como metadatos adicionales a la descripción de la obra.

- b. Además, en esta etapa es necesario establecer el esquema de metadatos adecuado a la publicación y tipo de información que incluye, por lo que es necesario contemplar la descripción de los criterios de calidad de software aplicados como parte de la documentación. Al respecto es importante señalar, que el tipo de información o correlación clara entre la asociación de cada elemento de información y su adecuada especificación en un esquema no es tema de esta investigación, sin embargo, en muchos sentidos los metadatos corresponden a características técnicas.
- c. También es necesario considerar personal encargado de esta etapa del proceso, tanto operativo como para la vigilancia en general de la calidad del software, pues dependiendo del tipo de publicación es posible que las verificaciones requieran recursos especiales tanto tecnológicos o laboratorios de prueba.
- d. Por lo anterior, asegurar la calidad de la publicación como software conlleva a establecer actividades que incluso definan y evalúen la idoneidad de los procesos de editoriales de tal forma que sea posible crear la confianza de que estos son los adecuados para lo que se desea hacer.

Así entonces, nuevamente se tienen cambios en la etapa de *producción* con la incorporación del rol de *programador* para *integrar funcionalidad* en la publicación cuando es necesario, además de la participación del *editor digital* para dar *seguimiento* en esta etapa del proceso, ver Figura 32.

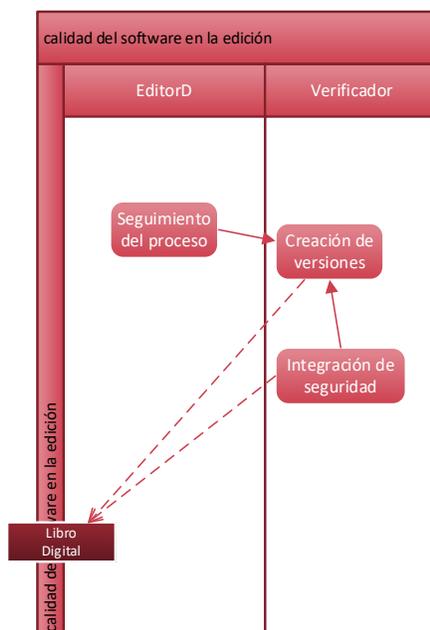
Figura 32. Cambios en la unidad de trabajo producción del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

Además, la etapa de *programación* cambia de nombre *control de calidad del software*, y se incorpora nuevamente al *editor digital*, para continuar con el *seguimiento* y deriva la responsabilidad al rol de *verificador*, encargado de ejecutar las acciones operativas para revisar la *composición* de la publicación, las *versiones* generadas e integrar elementos de *seguridad*, ver Figura 33, todo esto basado en la arquitectura definida y conformada en la etapa de *producción* y establecido desde la etapa de *control editorial*.

Figura 33. Cambios en la unidad de trabajo control de calidad del software en la edición.



Fuente: La autora diciembre 2019.

Respecto a las habilidades que el *verificador* y el *programador* deberían desarrollar para un mejor desempeño, con base en lo especificado en ESSENCE, es necesario que el programador cumpla con aptitudes que le permitan diseñar o construir software siguiendo estándares y normas establecidos por el equipo y respecto al verificador o especialista en pruebas, debe ser capaz de confirmar que el software cumple con los requerimientos establecidos, que la publicación es un producto usable y cubre los criterios establecidos por el equipo para determinar que está lista para ser distribuida, lo cual como ya vimos está definido en los atributos de calidad que ya se mencionaron.

#### 2.5.4 ETAPA DE: DISTRIBUCIÓN

La última etapa del proceso editorial propuesto está relacionada con la *distribución* de la publicación, que en estricto sentido es parte de los alcances definidos para la obra, por lo que es importante que la editorial u organización compruebe que hace la entrega o envío de todas las versiones, en las plataformas o medios establecidos, bajo los criterios

determinados como parte de la comercialización, difusión o resguardo de la publicación, ver Figura 34.

Figura 34. Distribución. Unidad de trabajo del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

En realidad, en el proceso definido hasta este momento sólo se contempla dentro de la *distribución* el definir los medios en los que se genera la publicación, sin embargo, vamos a ver que no es lo único a considerar, por lo que vale la pena revisar lo que ESSENCE define como parte de *las cosas que siempre se hacen*, respecto al *despliegue del sistema*, que se refiere a tomar un producto probado y colocarlo como disponible fuera del ambiente de desarrollo. En donde, además, se tiene un conjunto de elementos que nos ayudan a determinar que el producto está listo para su despliegue, lo cual implica que:

- Se cuenta con la documentación de instalación y uso del producto.
- Los representantes de los interesados aceptan que el producto está adecuado a lo propuesto.
- Los representantes de los interesados quieren usar u operar el producto.
- Es posible soportar la operación del software en un ambiente productivo.

Todo esto de alguna forma se trabajó en la etapa de *control de calidad del software* durante la edición, que como ya se vio abarca elementos que van más allá del contenido, sin embargo, el último punto además implica dar soporte al uso y para eso ESSENCE contempla dos acciones importantes:

- Empaquetar el producto para entregarlo o liberarlo en un entorno productivo.
- Hacer que el producto comience a usarse.

Además, en ESSENCE dentro de *las cosas que siempre se hacen*, también define elementos para el *Sistema en operación* que se refiere a soportar el uso del producto de software en un ambiente vivo o productivo y que implica:

- Mantener ciertos niveles de servicio
- Dar soporte a los interesados en el software
- Dar soporte o apoyo a quienes despliegan, usan o mantienen el uso del software

Esto implica un cambio importante para el nuevo proceso editorial, pues a diferencia del proceso tradicional, en donde una vez terminada la publicación el editor dejaba el producto completamente en manos de su usuario, concluyendo ahí su participación, para la publicación digital no va a suceder así, pues ahora el editor debería apoyar a los lectores otorgando ciertos niveles de servicio, por ejemplo, versiones para diversas plataformas de sistemas operativos o dispositivos, apoyo en la instalación o entregar manuales, de ahí que implica asumir nuevas responsabilidades con los lectores.

Además, el soporte contempla a los lectores y a todos los participantes en la cadena de valor del libro o ciclos de vida, como distribuidores, librerías y como ya se mencionó, también contempla a las instituciones que dan acceso a los materiales como bibliotecas o entidades receptoras de Depósito Legal encargadas del resguardo y preservación de las obras, para ayudar incluso en asuntos relacionados con la preservación de la éstas.

Sin embargo, en esta etapa también es necesario cerrar el proceso editorial, pues por lo general la obra ya no sufre cambios, a menos que ésta tenga como parte de sus alcances la renovación de funcionalidades de forma periódica, pero en lo general el proceso debe concluirse y establecer cuál es el alcance del soporte proporcionado a cada tipo de usuario de la obra, desde los lectores hasta las instituciones depositarias.

Desde luego que el *editor digital*, va a dar seguimiento al proceso editorial y se encargará de que todo concluya como se estableció, sin embargo cuando la publicación es enviada a un ambiente o medio de distribución, comercial o no, debería existir un nuevo rol que se encargue de atender las actividades de soporte y mantenimiento de la publicación, pues además es quién en algún momento recopila nuevas necesidades y cambios que generen la producción de nuevas manifestaciones de la obra o nuevas versiones, derivadas de la identificación de nuevos requerimientos, más aún cuando se trate de obras que por sus alcances tienen una renovación continua de funcionalidades, con lo cual entran en un proceso editorial nuevamente o de forma periódica.

De ahí que, para esta etapa, es necesario considerar las siguientes acciones derivadas de la Ingeniería de Software:

**1. En cuanto a la liberación o distribución de la obra:**

- a. Es necesario establecer medios que apoyen o soporten el uso de la publicación, desde manuales de usuario hasta servicios en línea que orienten o resuelvan inquietudes de los lectores o de instituciones encargadas del Depósito Legal.
- b. Para las publicaciones digitales, el contar con personal dedicado al soporte de usuarios es parte de las actividades que deben planearse desde el proceso editorial, por lo que mientras más sofisticada es la publicación, es más probable requerir de personal más especializado para apoyar estas tareas que al final también son parte de las necesidades de los lectores.
- c. La usabilidad es un atributo importante para todo producto digital, pues mientras más se trabaje en productos intuitivos y fáciles de usar, menor será la incidencia de problemas en los lectores y los servicios de soporte que se requieran.
- d. Es necesario considerar que como parte de los interesados en una publicación se encuentran las instituciones receptoras del Depósito Legal, por lo que, dentro de los requisitos generales a satisfacer, están los que permitan la preservación a largo plazo de la obra, de ahí que parte de las labores de soporte se dirijan en este sentido.

**2. En cuanto al cierre del proceso editorial:**

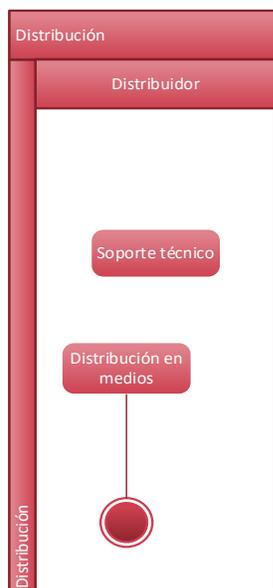
- a. Todo proceso editorial digital debe tener un momento en el que se concluye y libera o distribuye una obra, por lo que dentro de los alcances de la publicación es necesario establecer por escrito o de forma muy clara el momento en que la obra se dará por terminada o en su caso liberada al público y deja el proceso de edición para pasar a un ambiente de distribución. Esto no significa que la publicación pueda entrar en un proceso de edición de nuevo en corto, mediano o largo plazo, lo importante es señalar que existen criterios que determinan el momento en que la obra se hace pública.
- b. El cierre del proceso editorial digital implica que la publicación se encuentra empaquetada con las manifestaciones e ítems generados de acuerdo con lo establecido en los alcances, en el lugar o medios de distribución señalados por la editorial.

Por lo tanto, la etapa de distribución forma parte del proceso editorial de libro digital, en tanto que es necesario contemplar que la publicación una vez concluida y verificada, debe encontrarse en un ambiente en el que esté vigilada por un *distribuidor* quien se encarga de hacer el envío de la obra a todos los medios de distribución con base en lo establecido en su proceso editorial y además brinda el soporte necesario a los usuarios sean estos comercializadores, lectores e incluso instituciones receptoras de Depósito Legal, ver Figura 35.

A pesar de que el proceso editorial de una obra concluye con la distribución, el soporte técnico sigue siendo responsabilidad del editor, sin embargo, éste no es parte del proceso editorial, pero si es un elemento que debe considerarse dentro de los alcances de la

publicación, para establecer el tipo de servicio o soporte que se otorgará al usuario al momento de distribuir la obra, considerando como ya se dijo antes, a todos los involucrados en el ciclo de vida del libro.

Figura 35. Cambios en unidad de trabajo distribución del proceso editorial propuesto.



Fuente: La autora diciembre 2019.

### 2.5.5 LA ESPECIFICACIÓN DE UN PROCESO DE EDICIÓN FORTALECIDO POR LA ESENCIA DE LA INGENIERÍA DE SOFTWARE.

Con todo lo anterior, se generó un proceso fortalecido con cada una de las etapas que aquí se fueron describiendo, ver Figura 36, orientado a generar un producto que en este caso llamamos libro digital, pero que en esencia como todo producto de software va a requerir no sólo de una construcción ordenada y controlada, además va a imponer nuevas opciones de servicio para los editores y características de calidad en la publicación que van más allá del contenido.

Además, se puede decir que el proceso editorial definido en este trabajo de investigación tiene factores que van a afectar el sistema organizacional o entorno en el cual se desarrolla y que impactan en todas las etapas (*control editorial, producción, control de calidad del software y distribución*):

- a. **El software:** que se refiere a las aplicaciones que son generadas como un producto o publicación y las utilizadas como herramientas durante los procesos de edición, pues ambas deben ser evaluadas desde una perspectiva de preservación.

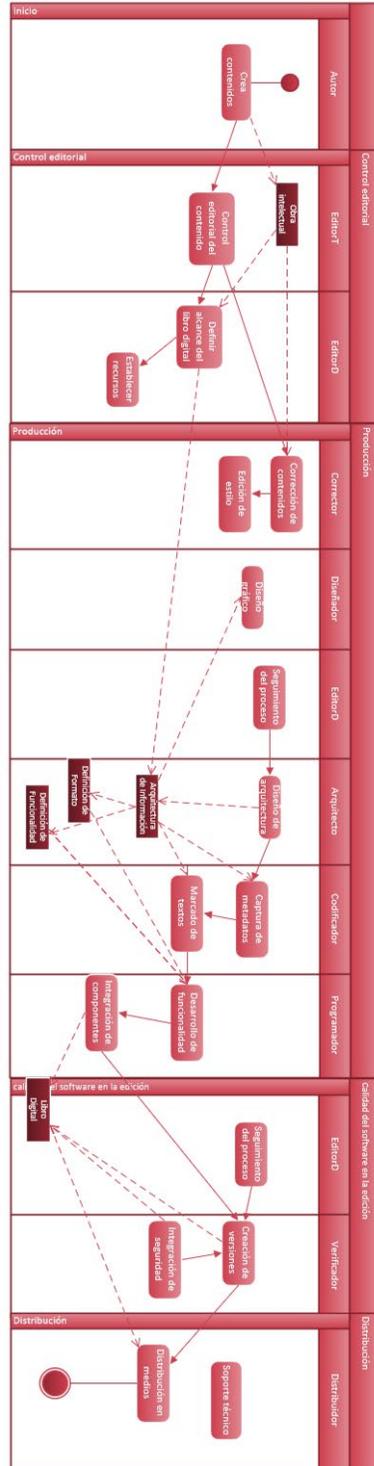
- b. **Los procedimientos:** Se trata de las acciones encaminadas a alcanzar un objetivo concreto y que en su conjunto van a formar el proceso editorial, por lo que es necesario contar con la especificación clara y por escrito de los pasos incluidos en cada etapa del proceso.
- c. **Las reglas de negocio:** Deben conocerse, entenderse, describirse, validarse, implementarse, probarse y asumirse todos los procedimientos que formaran parte del proceso editorial de un libro. Esto contempla desde la integración de recursos humanos, hasta las posibilidades de servicio que formarán parte del soporte de las publicaciones una vez liberadas, además de los controles de calidad que se requieren para generar productos que abarquen las dos perspectivas que tiene la publicación, el contenido intelectual y la calidad del software.
- d. **Los recursos humanos:** Se refiere al personal calificado con nuevas habilidades que va a aportar conocimiento para determinadas acciones, de acuerdo con las capacidades y las necesidades o alcances establecidos para la publicación. De ahí que el proceso editorial de libros digitales va a requerir de nuevos perfiles profesionales para el manejo de nuevos estándares o elementos de calidad.
- e. **La documentación asociada:** Es claro que los datos generados durante el proceso editorial forman parte importante de la historia de una publicación digital en donde además de los datos descriptivos, están los técnicos que como ya se comentó se refieren a los atributos de calidad del software.

De ahí que existen elementos clave que permiten apoyar la calidad de la publicación digital:

- a. **Establecer alcance y composición de la publicación.** Que incluye la especificación de la estructura de la publicación y la Identificación de recursos a utilizar de acuerdo con el alcance deseado.
- b. **Definir criterios básicos de calidad.** Que contempla el uso de estándares y la reutilización recursos o prácticas generadas previamente en la organización o por la comunidad, de acuerdo con el alcance deseado, que indiscutiblemente debe incluir la integración de metadatos.
- c. **Trabajo colaborativo de todos los involucrados.** Que contempla solicitar o incluir el apoyo de expertos para cada actividad nueva y desde luego contempla los pasos ya conocidos para los procesos editoriales tradicionales, considerando lo necesario incluso, cuando se trate de libros impresos que se publican también en formato digital.
- d. **Verificación y control del proceso de producción y de la publicación.** Que incluye la documentación del proceso y el seguimiento para garantizar la calidad esperada en el resultado.

En conclusión, es necesario establecer una estructura lógica y uniforme de estos elementos, en la búsqueda de un mejor resultado para la definición de un modelo que a partir de un proceso para la edición de libros digitales considere también los factores que giran alrededor de la edición, mencionados al inicio de este capítulo como ciclo de vida y cadena de valor, lo cual, se analizará en el siguiente capítulo.

Figura 36. Proceso editorial con adecuaciones derivadas de aplicar IS.



Fuente: La autora diciembre 2019.

## CAPÍTULO III. DESARROLLO DEL MODELO.

Un modelo se usa para captar y enumerar exhaustivamente los requisitos y el dominio de conocimiento, de forma que todos los implicados puedan entenderlos y estar de acuerdo con ellos (Rumbaugh et al., 2007, p. 11).

Este capítulo inicia con la definición del concepto de modelo, con lo cual se establecerá la forma en que se presentará el desarrollo del objetivo principal de esta investigación y a partir de eso, se irán desatando los elementos que fundamentan *el modelo de integración de metadatos para preservación en la edición del libro digital*, que de forma general se desprende de la visión del proceso de edición creado en el capítulo anterior y la comparación de elementos derivados de los modelos FRBR y OAIS, el ciclo de vida del libro definido por Darnton, la cadena de valor del libro y los atributos de calidad de software, pues son los factores y conceptos que giran alrededor del libro digital y como ya se analizó en los capítulos anteriores respaldan la necesidad de incluir metadatos para preservación en la edición.

### 3.1. ¿QUÉ ES UN MODELO?

La palabra modelo de acuerdo con el diccionario de la Real Academia Española puede entenderse desde varias perspectivas, ya sea como una persona u objeto, que es copiado por un artista o arquetipo que puede ser imitado o reproducido y también, como un esquema teórico de un sistema o realidad compleja (RAE & ASALE, 2020).

El filósofo Alejandro Cassini (Cassini, 2016) señala que el concepto de modelo adquirió un papel preponderante para la filosofía de la ciencia, desde finales del siglo XX y hasta ahora, colocando en el centro de atención el concepto de modelo más que el de teoría, con lo cual, hay quienes han categorizado el desarrollo de la filosofía de la ciencia como “modelística” (Cassini, 2016), a partir de 1970. Además, marca las siguientes características para distinguir a los modelos de las teorías, a pesar de que estas últimas, siguen siendo esenciales en la ciencia:

- a) Los modelos suelen tener un ámbito de aplicación sumamente restringido y acotado.
- b) Tienen un carácter híbrido y están formados por hipótesis que pertenecen a diferentes teorías, además de incorporar datos empíricos de diferentes niveles.
- c) Pueden tener una vida efímera.
- d) Presentan un cierto grado de idealización.
- e) Y tienden a emplear múltiples modelos, para dar cuenta de un mismo dominio de fenómenos.

Así pues, un modelo teórico puede estar definido como la “construcción física de un caso dado o incluso un conjunto abstracto de ideas” (Achinstein, 1967) que parte de una construcción mental para permitir una aproximación a la realidad de un fenómeno, distinguiendo sus características para facilitar su comprensión.

La creación o definición de modelos es una estrategia muy usada en diversas disciplinas y una práctica común, de ahí que estos pueden conceptualizarse como mapas, maquetas, íconos, prototipos, sistemas de ecuaciones y simulaciones computacionales (Cassini, 2016). En donde el modelo puede verse como un objeto concreto, por ejemplo, una maqueta, que sirve de aproximación a lo que se pretende construir o una ecuación, en la que el modelo es abstracto y busca hacer una predicción teórica o ayudar en la enseñanza de la ciencia. De ahí que, “la finalidad con la que se construyen los modelos depende de los intereses de los usuarios de tales modelos” (Cassini, 2016).

Sin embargo, Achinstein considera que cualquier modelo teórico tiene al menos cuatro características (Achinstein, 1967):

1. Consiste en un conjunto de supuestos sobre algún objeto o sistema.
2. Describe un tipo de objeto o sistema atribuyéndole lo que puede llamarse una estructura interior, composición o mecanismo, con referencia al cual se explicarán varias propiedades exhibidas por ese objeto o sistema
3. Se utiliza como una aproximación útil para ciertos propósitos.
4. Frecuentemente se formula, desarrolla y designa sobre la base de una analogía entre el objeto o sistema descrito en el modelo y algunos objetos o sistemas diferentes. Esto entraña una comparación por la cual se advierten propiedades y principios en cierto sentido similares.

Por lo anterior, para entender este concepto es necesario hablar de lo que es un sistema, pues Achinstein también asegura que:

el estudio científico de un modelo aspira a elaborar una teoría del sistema, es decir, un conjunto de enunciados, ecuaciones, formulas, esquemas, etc. Que permitan describir adecuadamente el funcionamiento presente del sistema, así como explicar lo ocurrido en el pasado y predecir lo que pasará en dicho sistema en el futuro (Achinstein, 1967)

De ahí que, un *sistema* puede considerarse como “una colección de objetos o entidades que interactúan entre sí para alcanzar un cierto objetivo” (Guasch, 2009). Cabe aclarar que el estado u objetivo que puede alcanzar el sistema depende de los objetos y entidades que interactúan en determinado momento dentro del sistema. Y el *estado de un sistema* puede ser definido como:

el conjunto mínimo de variables necesarias para caracterizar o describir todos aquellos aspectos de interés del sistema en un cierto instante de tiempo (Guasch, 2009).

Al proceso para abstraer las características de un sistema que permita generar un modelo se le llama *modelado* y este, generalmente requiere de un lenguaje o conjunto de símbolos, que permiten expresar de forma clara lo que representa el modelo.

El filósofo Alejandro Cassini (Cassini, 2016) refiere que, por la diversidad de significados y usos, es muy difícil e incluso prematuro intentar clasificar los tipos de modelos que se emplean en la ciencia y por eso, es mejor sólo caracterizar los más usados:

- Modelos matemáticos, empleados en la lógica matemática como *teoría de modelos*, en donde, un modelo es una interpretación de una teoría formal en el cual, todas las fórmulas de dicha teoría resultan verdaderas.
- Modelos en ciencias naturales y sociales, incluye desde íconos, prototipos, maquetas, mapas, diagramas, etc. que se emplean para hacer una representación aproximada de un fenómeno determinado.
  - Modelos de datos. Representación idealizada generada mediante la aplicación de instrumentos estadísticos para hacer mediciones repetidas a los datos crudos.
  - Modelos con una función heurística, predictiva, didáctica o exploratoria. Permiten el acceso a fenómenos poco conocidos o que no resultan tratables con los recursos del conocimiento vigente, pueden ayudar a predecir la ocurrencia de fenómenos o pueden ser útiles para introducir a los estudiantes en temas complejos.
  - Modelos con función mecánica o explicativa. Buscan explicar o presentar los mecanismos causales para la producción de un fenómeno y en algunos casos las explicaciones son no causales.

Además, Cassini (Cassini, 2016) señala que es importante entender lo complicado que es definir por completo cuál es la relación que existe entre los usuarios y sus modelos<sup>169</sup>, sin embargo, se puede asegurar que:

- Los modelos científicos tienen una diversidad de usos y funciones que parecen ser irreductibles.
- Los modelos se construyen para resolver un problema determinado en un cierto dominio de fenómenos, aunque frecuentemente tienen aplicaciones en dominios de fenómenos no previstos.
- Con adaptaciones, son incluso capaces de migrar de una ciencia o disciplina a otras muy diferentes y aparentemente alejadas entre sí, por lo que es posible usarlos como plantillas o estructuras formales o computacionales.
- Los modelos siempre se construyen teniendo en cuenta los intereses de los usuarios y están sujetos a cambios cuando estos intereses se modifican o se transforman.

Así pues, se puede afirmar que, para la explicación u exploración de acciones involucradas en un proceso o fenómeno, es posible utilizar modelos que permita describir su ocurrencia

---

<sup>169</sup> Al estudio de la relación entre los modelos y sus usuarios, se le llama pragmática de los modelos

o simplemente esbozar su composición e interacción para comprender su comportamiento como un sistema.

De ahí que, la concepción de modelo, desde el punto de vista de la automatización de procesos, en donde el modelo es la descripción de características internas de un sistema al que se quiere imitar su comportamiento (Guasch, 2009), es una práctica muy utilizada en diversas áreas de conocimiento, incluida la Ingeniería de Software (IS) en donde:

Un modelo se usa para captar y enumerar exhaustivamente los requisitos y el dominio de conocimiento, de forma que todos los implicados puedan entenderlos y estar de acuerdo con ellos (Rumbaugh et al., 2007, p. 11) .

Así pues, la IS ha definido lenguajes de modelado para llevar a cabo el análisis de procesos, sin embargo, uno de los más utilizados para el diseño de software es el Lenguaje Unificado de Modelado, UML por sus siglas en inglés, con lo cual es posible analizar las implicaciones tecnológicas de un proyecto de IS o creación de software, antes de invertir grandes cantidades de recursos, ya que:

UML capta la información sobre la estructura estática y el comportamiento dinámico de un sistema (Rumbaugh et al., 2007, p. 3).

En UML, un sistema es una colección de objetos que interactúan para ejecutar un trabajo, que beneficia a un usuario. La parte estática permite identificar a los objetos o elementos importantes en un sistema y la relaciones que existen entre ellos, y la parte dinámica, se refiere a la historia de los objetos y su comunicación con el resto, para cumplir ciertos objetivos en un momento dado (Rumbaugh et al., 2007).

El modelado, desde este punto de vista, considera los siguientes aspectos (Rumbaugh et al., 2007):

- **Abstracción frente a detalle.** El modelo capta elementos esenciales de un sistema y omite otros, en donde lo esencial dependerá del propósito del modelo.
- **Especificación frente a implementación.** Es importante definir el qué, antes de especificar el cómo.
- **Descripción frente a instancia.** Los modelos son principalmente una descripción de instancias que aparecen en tiempo de ejecución.
- **Variaciones en la interpretación.** El modelo puede tener varias interpretaciones, por lo que, es adecuado marcar *puntos de variación semántica* y asignar a cada interpretación un nombre.

Y desde la perspectiva de UML los modelos van a tener tres aspectos importantes: *semántica*, *presentación visual* y *contexto*. La *semántica* captura construcciones lógicas, por lo tanto, cada elemento del modelo tiene un significado, de ahí que requiere de una estructura sintáctica basada en reglas las cuales se definen en el lenguaje de modelado. La *presentación visual*, es la manera de presentar la información semántica comprensible a las

personas. Y el *contexto*, es importante ya que por lo general los sistemas, siempre se utilizan dentro de un ámbito más grande o una organización y es necesario que el modelo considere información referente al entorno (Rumbaugh et al., 2007).

De ahí que la Ingeniería de Software cuenta con herramientas que permiten comprender la semántica que gira alrededor de un sistema y los mecanismos para comunicar o transmitir esto, de tal forma que las técnicas de modelado y procesos de construcción empleados están encaminados a generar resultados adecuados a las necesidades identificadas en un sistema, con un software como resultado.

### 3.2. BASES DEL MODELO DESCRITO EN ESTE TRABAJO DE INVESTIGACIÓN

El *modelo de integración de metadatos para preservación en la edición del libro digital* que aquí se define, de acuerdo con lo referido por Cassini, es un modelo con función mecánica o explicativa aplicable en las ciencias sociales, con el cual se presentan los mecanismos que permiten la integración de metadatos para preservación en el fenómeno conocido como edición del libro digital y se basa en lo que Achinstein menciona respecto a las características de un modelo, por lo que cumple con las siguientes características:

1. Establecer un conjunto de supuestos respecto al libro digital, su edición y metadatos para preservación.
2. Describir el sistema que gira alrededor del libro digital, de tal forma que se implante el sistema organizacional que lo respalda y exhiba las propiedades básicas del libro respecto a los metadatos para preservación.
3. Sirve como una aproximación de lo que implica la edición del libro digital para la integración de metadatos para preservación.
4. El modelo parte de las perspectivas establecidas en la calidad del libro digital que abarca su contenido intelectual y su estructura y composición digital contenidos como metadatos para preservación.

También sabemos que los modelos incluyen desde íconos, prototipos, maquetas, mapas, diagramas, etc. que se emplean para hacer una representación aproximada de un fenómeno determinado, el cual se conoce como lenguaje de modelado.

Así pues, para este trabajo de investigación se usará UML como lenguaje de modelado, que como ya se explicó anteriormente, permite describir la semántica, presentación visual y contexto, por lo que es utilizado en Ingeniería de Software para el modelado de procesos partiendo de las siguientes premisas (Rumbaugh et al., 2007):

- **Especificación frente a implementación.** Es importante definir el qué, antes de especificar el cómo.
- **Abstracción frente a detalle.** El modelo capta elementos esenciales de un sistema y omite otros, en donde lo esencial dependerá del propósito del modelo.

De ahí que, el modelo que aquí se presenta se utilizará para “captar y enumerar los requisitos y el dominio de conocimiento” (Rumbaugh et al., 2007, p. 11) en el que se desarrolla la integración de metadatos en la edición del libro digital, con la finalidad de que “todos los implicados puedan entenderlos y estar de acuerdo con ellos” (Rumbaugh et al., 2007, p. 11).

Además, el modelado se basa en tres aspectos: la semántica, en donde se capturan construcciones lógicas y cada elemento tiene un significado; la presentación visual, por lo que la información que se modela es comprensible para las personas, pues se utiliza un lenguaje en donde cada elemento tiene un significado y respeta una estructura sintáctica basada en las reglas definidas por UML; y por último el contexto, que describirá la información correspondiente al entorno que deberá generarse, para propiciar la integración de metadatos para preservación en la edición del libro digital.

Cabe resaltar, que estos tres aspectos se utilizan dentro del modelado al mismo tiempo, durante la definición del modelo, por lo que, en adelante, se utilizarán como herramientas para la descripción y explicación del desarrollo de este trabajo de investigación.

### 3.3. ESPECIFICACIÓN FRENTE A IMPLEMENTACIÓN.

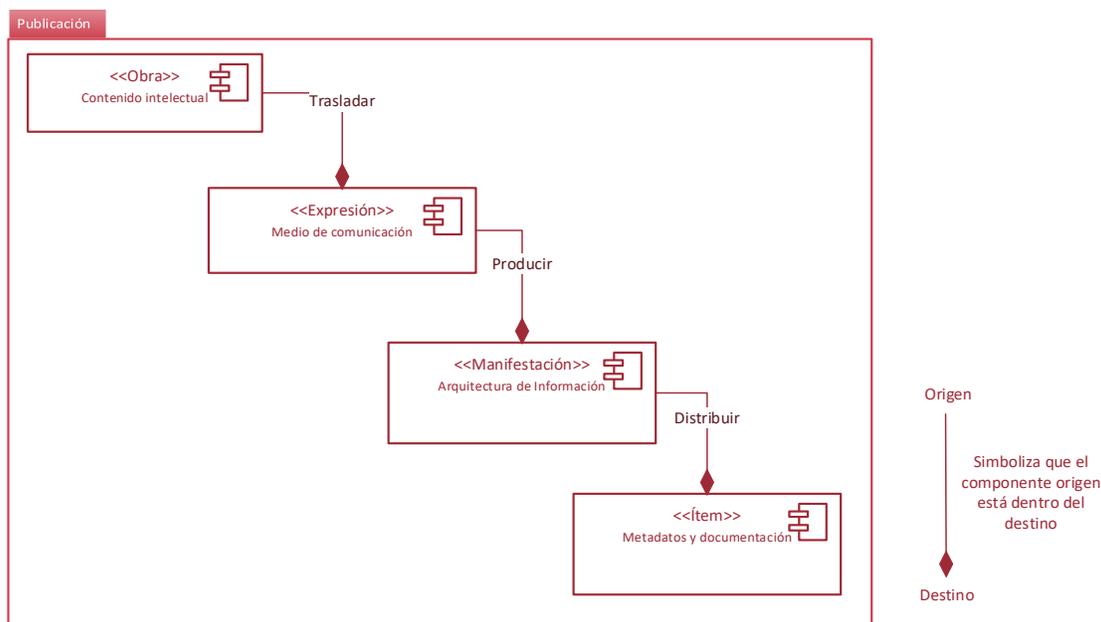
El modelo que aquí se presenta, como ya se mencionó, comienza al definir el qué, antes de especificar el cómo y en ese sentido, lo primero será la revisión del libro digital como publicación a partir de sus características definidas con el modelo conceptual FRBR<sup>170</sup>, que como ya se vio en otro capítulo, explica la composición del universo bibliográfico a partir de cuatro entidades: *Obra*, *Expresión*, *Manifestación* e *Ítem* y que al insertarlo en el modelo que aquí se propone, se representará con un diagrama UML<sup>171</sup> de componentes, para simbolizar al libro o publicación y las relaciones que se generan a partir de su edición, ver Figura 37.

Así entonces, se tiene a la *obra* caracterizada por su **contenido intelectual**, que se va a “*trasladar*” a una *expresión*, en donde gracias al empleo de un **lenguaje de comunicación** es posible expresar el contenido, sin embargo para formalizar la difusión de la expresión, se requiere “*producir*” la *manifestación*, lo cual implica un proceso editorial, en el cual se crea el contenedor adecuado y por lo tanto, como ya se vio en el capítulo anterior, se establece una **arquitectura de información** a partir del tipo de *expresión* y el alcance de la *manifestación*, que finalmente se va a “*distribuir*” como un *ítem*, que como ya se vio en el proceso editorial propuesto, ésta unidad física, al momento de distribuirse debería contener además de la obra intelectual, los **metadatos y documentación** de la publicación, necesaria para ser utilizada.

<sup>170</sup> Requerimientos Funcionales para Registros Bibliográficos (<https://www.loc.gov/catdir/cpsa/Que-es-FRBR.pdf>)

<sup>171</sup> Los diagramas de componentes UML (Lenguaje Unificado de Modelado) son usados para representar las partes que van a conformar un paquete de software y sus dependencias

Figura 37. Modelo, nivel 1, diagrama de componentes para especificar a la publicación.



Fuente: La autora diciembre 2019.

Cabe señalar que los **metadatos y documentación** del ítem, en esencia deberían permitir entender a las entidades contenidas, dado que estas se encuentran inmersas, una dentro de la otra, derivado de relaciones de transformación o acciones que permiten perfeccionar una obra para ser accesible como publicación, que, en esencia es lo que representa este diagrama del nivel 1 del modelo y con lo cual se identifica la composición del primer elemento del modelo, **la publicación**. Es importante mencionar que, para efectos de este trabajo de investigación, al referirnos a **la publicación**, se hace referencia al libro digital.

Lo siguiente ahora, es identificar a las entidades con las que la publicación se relaciona y para eso nuevamente retomaremos FRBR, que como ya vimos en el diagrama del grupo2 considera a dos entidades del mundo real: *persona* y *entidad corporativa*, que en esencia simbolizan la interacción de un individuo o una institución con la publicación.

Para esto también es necesario identificar quiénes participan en el entorno que gira alrededor de un libro digital por lo que es conveniente recordar lo planteado en el capítulo anterior, por Darnton cuando habla de un “circuito de comunicación” para el libro (Robert Darnton editado por Finkelstein et al., 2006, p. 12), que él define como ciclo de vida con el cual se establecen participantes de dicho circuito que han acompañado al libro desde hace varios siglos: autor, editor, impresor, distribuidor, librero y lector que incluye a las bibliotecas, que como ya se dijo en su momento, hay algunas que funge como entidades depositarias de la publicación.

De ahí que, para el modelo que aquí se describe cualquiera de las dos entidades: *persona* y *entidad corporativa* podría representar a los participantes mencionados por Darnton, pues

estos pueden ser quienes crean la obra, realizan la expresión, producen la manifestación y en su momento poseen el ítem, ya sea porque lo adquieren como lectores o como instituciones encargadas del resguardo o protección del material.

A partir de esta conceptualización, es posible decir que:

- una *obra* puede ser *trasladada* a una *expresión* por una o varias personas
- una *expresión* puede ser *producida* como una *manifestación* por una o varias personas
- una *manifestación* puede ser *distribuida* como un *ítem* a una o varias personas

Destaquemos pues, que existen al menos tres grandes tipos de personas u organizaciones de personas, relacionadas con la publicación. Las que intervienen en la creación de la obra y expresión, que para fines de esta investigación llamaremos *Ente-Autor*, pues en esencia son quienes construyen las ideas o conceptos que definen el trabajo intelectual y lo plasman en un medio de comunicación que puede ser escrito, visual o auditivo para ser expresado.

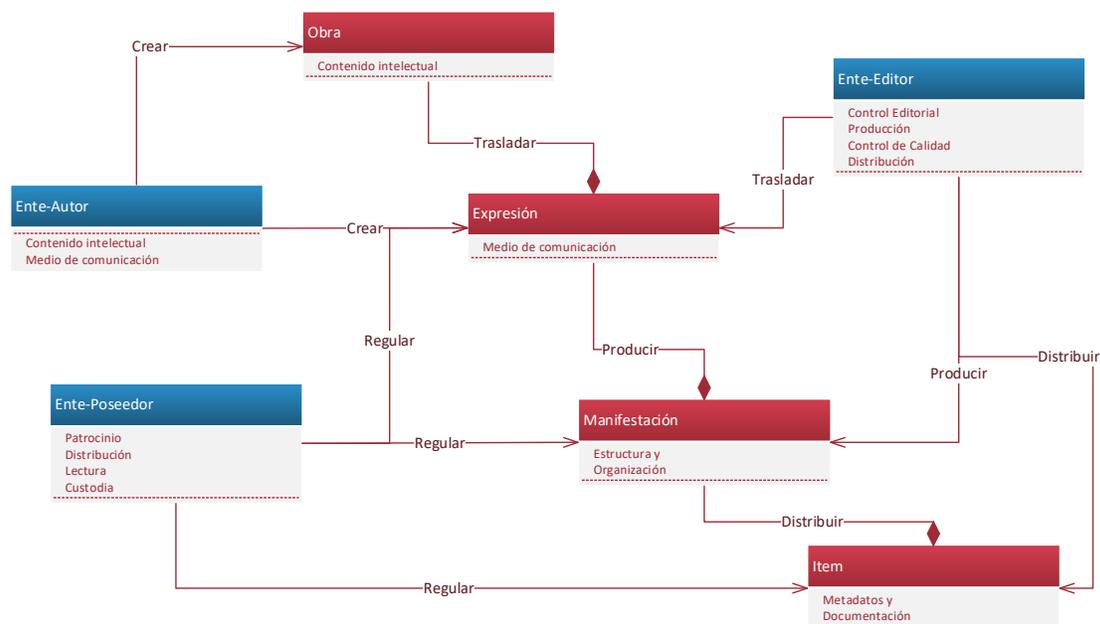
También existen los individuos que intervienen en la producción de la manifestación, que para fines de esta investigación llamaremos *Ente-Editor*, que como ya se vio en el proceso editorial propuesto en el capítulo anterior, incluye muchos tipos de participantes, que van desde quienes lo controlan hasta quienes se encargan del soporte de la publicación.

La tercera agrupación se refiere a quienes poseen un ítem, ya sea porque lo distribuye o vende, lo cuida o protege o lo adquiere para uso individual, de ahí que en esta investigación lo llamaremos *Ente-Poseedor*.

Así pues, para expresar las relaciones entre la publicación y estos elementos, utilizaremos un diagrama UML de clases, con el cual se representará objetos del mundo real y conceptos para establecer la forma en que se relacionan y los atributos que los caracterizan. De ahí que, trasladaremos los componentes de la publicación y los participantes que se han descrito, a clases o conceptos representados en el diagrama nivel 2 del modelo que muestra a los involucrados en la producción de la publicación, ver Figura 38.

Es claro que la clase *Ente-Autor* siempre va a referirse al o los creadores de obra, sin embargo, también puede incluir la creación de la expresión. Aquí es conveniente señalar que, dadas las nuevas maneras de expresión de una publicación digital, esta clase de objeto puede agrupar a escritores, traductores, locutores, cineastas, animadores, ilustradores y programadores. Pues la obra desde un punto de vista amplio puede referirse a una aplicación interactiva, una base de datos, sitio web, videojuego en donde los creadores son más de una persona, de ahí el término publicación, sin embargo, para fines de este trabajo de investigación, como ya se dijo la publicación se refiere al libro digital.

Figura 38. Modelo, nivel 2, diagrama de clases para representar involucrados en la producción de la publicación.



Fuente: La autora diciembre 2019.

Sin embargo, en este trabajo no se pretende disertar en cuanto al concepto de autor, por lo que aquí, sólo se establece a esta entidad como representante de todos los participantes en la conceptualización de la obra, sin que esto implique un mayor alcance, pues es claro que tradicionalmente este concepto aún es un punto de discusión. Así que, sólo para fines prácticos, entenderemos a la clase *Ente-Autor* como una representación de los individuos u organismos capaces de conceptualizan una obra y que, en mayor o menor medida, participarán en el traslado de ésta, a una expresión.

El *Ente-Editor* dentro del modelo va a agrupar a todos los participantes en la edición, desde el director editorial, patrocinadores, redactores, correctores, revisores, traductores, impresores, etc., y los nuevos participantes en la producción digital como el editor digital, codificadores, arquitectos, programadores, etc. Pues se trata de la entidad que va a representar a todos los individuos u organismos encargados de construir la expresión. Que como se puede ver en el mismo diagrama, trabaja a la par con el *Ente-Autor*, sin embargo, el *Ente-Editor* es el responsable de llevar a cabo las cuatro etapas definidas en el proceso editorial propuesto en el capítulo anterior: *control editorial, producción, calidad del software en la edición y distribución*.

Cabe señalar que, establecer un *Ente-Editor* para este modelo no intenta cambiar la conceptualización teórica de lo que es la edición, pues sólo sirve para fines prácticos de este trabajo, como un concepto que figura dentro del modelo para agrupar los roles y etapas que participan en el proceso editorial del libro digital, sea cual fuera este.

La tercera clase señalada como *Ente-Poseedor* va a agrupar, para fines de este modelo, tres perspectivas de quienes se identifican como receptores de la publicación, basado en lo que señala Darnton y adaptado para fines del modelo:

1. Comercial, que incluye a quienes financian o patrocinan y distribuyen, es decir personas u organizaciones que aporta los recursos para la producción y/o distribución de la publicación, por lo que se consideran poseedores para comerciar con ella;
2. Consumo, son quienes compran o adquieren la publicación por interés personal conocidos generalmente como lectores;
3. Custodia, se refiere a las organizaciones que recaban la publicación para resguardo y procuran su cuidado a largo plazo, pues sirven de apoyo para el lector, ya que proporcionan alternativas de acceso diferentes a la compra.

Además, en este modelo, la entidad *Ente-Poseedor* establece una relación *reguladora* con la publicación, basada en el ciclo de vida del libro definido por Darnton, pues en el ciclo de vida señala que existen elementos que retroalimentan al libro y que regulan o influyen determinando los alcances de la publicación, por lo que al momento de su construcción es necesario tomar en consideración a esta entidad. Además, esta relación reguladora también se refuerza con lo planteado en este trabajo de investigación en el capítulo anterior, respecto a las prácticas de Ingeniería de Software, que recomiendan para la edición, considerar la cantidad de recursos con los que se cuenta, el público al que va dirigido y las necesidades a satisfacer de los interesados en el libro, que como ya se vio incluyen la preservación digital en las bibliotecas.

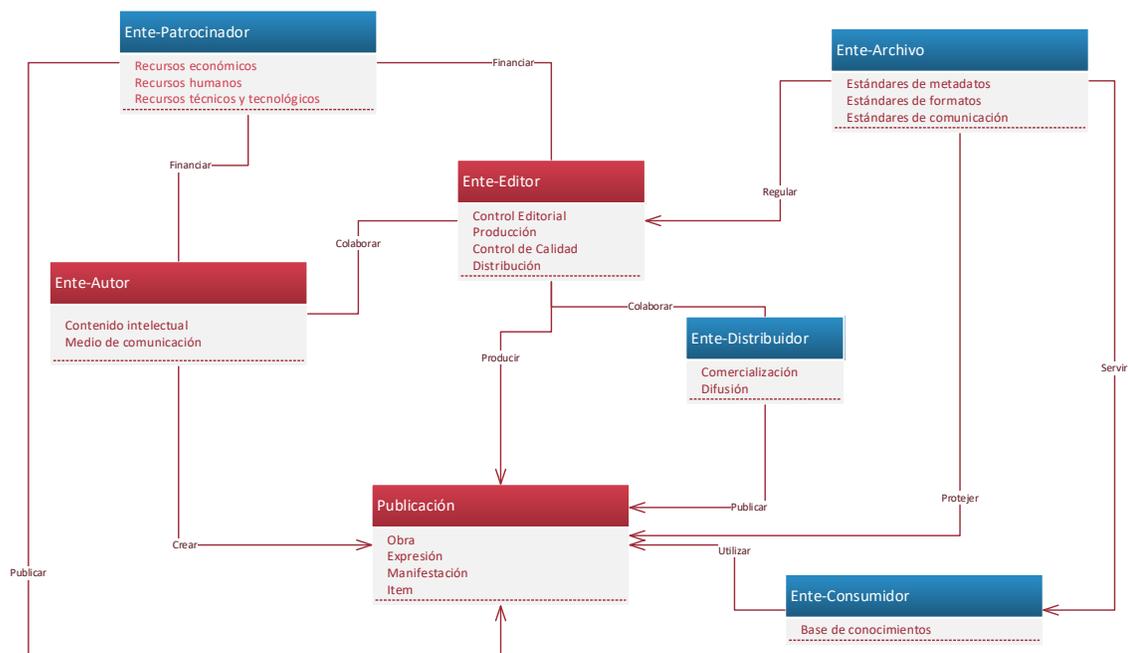
Por lo anterior y dada la importancia que tiene la preservación digital para este trabajo de investigación, se hará una disgregación de esta entidad, para identificar a cada uno de los involucrados y diferenciar el tipo de interacción que existe entre estos y la publicación.

En el diagrama nivel 3 del modelo, se van a representar a los involucrados, que, desde el punto de vista planteado aquí, intervienen en la definición de alcances de la publicación. Además, por cuestiones de optimización, en este nivel 3 del modelo, se concentrarán los componentes de la *publicación* en una sola clase.

De tal forma que las clases en las que se descompone *Ente-Poseedor*, definidas para fines de esta investigación son, ver Figura 39:

- *Ente-Patrocinador* va a *financiar* la *publicación*
- *Ente-Distribuidor* va a *distribuir* la *publicación*
- *Ente-Consumidor* va a *utilizar* la *publicación*
- *Ente-Archivo* va a *proteger* la *publicación*

Figura 39. Modelo, nivel 3, diagrama de clases para representar involucrados en el alcance de la publicación.



Fuente: La autora diciembre 2019.

La clase *Ente-Patrocinador*, se ha nombrado así para fines de esta investigación, con la intención de diferenciar a quién está encargado de aportar los recursos para la producción de la publicación e incluso también para la creación de la obra, por lo que siempre tendrá una relación de financiamiento con el *Ente-Editor* y el *Ente-Autor*, además, el *Ente-Patrocinador* es una figura abstracta que separa la parte financiera y de recursos, sin embargo en la vida real, puede ser la misma persona que aporta los recursos y lleva a cabo el proceso de edición, de ahí que su presencia en el modelo, sólo sirve para no perder de vista que la publicación siempre dependerá de los recursos asignados para la edición. Y, además, para simboliza la propiedad que adquiere el *Ente-Patrocinador* sobre la publicación, para comerciar con ella derivado de su aportación de recursos. Esta clase como el resto de las definidas en el modelo puede representar a un individuo o incluso a una organización.

La clase *Ente-Distribuidor* tiene una relación de colaboración con el *Ente-Editor* y además es el encargado de hacer público el producto editorial o publicación, ya sea con fines comerciales o de difusión.

Clase *Ente-Consumidor*, se refiere a quién utiliza la publicación, también denominada en el modelo OAIS como comunidad designada. Es importante recalcar que el consumidor dado que nos referimos a una publicación digital, va a incluir además del lector tradicional o personas, a una aplicación informática, como los robots de buscadores web e incluso

cosechadores. De ahí que, el consumidor es quien va a hacer uso del ítem y para entender el contenido de la obra requerirá información o bases de conocimientos que le permitan interpretar a la publicación. Por lo que en muchas ocasiones se servirá de la información que el *Ente-Archivo* pueda proporcionar al respecto de la publicación e incluso la publicación misma como ya se mencionó, deberá apoyar en esto, con los metadatos y documentación contenida.

*Ente-Archivo* o encargado de custodia recibe este nombre para estar en concordancia con el modelo OAIS, en donde un *Archivo* es la institución que asume la responsabilidad del resguardo o concentración de las publicaciones, para su preservación a largo plazo y, además, de acuerdo con el mismo modelo es encargada de “La negociación con quienes producen los objetos de información, para el ingreso apropiado al Archivo” (CCSDS, 2012, p. 3-1). De ahí que la relación con el *Ente-Editor* idealmente, es de regulador, pues es el encargado de establecer las pautas que van a encaminar al *Ente-Editor* para que genere una *Publicación* que pueda ser utilizada por el *Ente-Consumidor* y resguardada por el *Ente-Archivo*, de ahí que, su relación con la *Publicación* es de protección y con el *Ente-Consumidor* de servicio.

El *Ente-Archivo*, no se refiere a una única institución, pues es claro que para cada tipo de publicación existen interesados o responsables de su resguardo, de ahí que esta puede representar a cualquier persona, institución o conjunto de instituciones que deciden fungir como responsables del cuidado de un tipo de publicación, a largo plazo y que, para efector de esta investigación, es el que llevará a cabo los procesos de preservación digital.

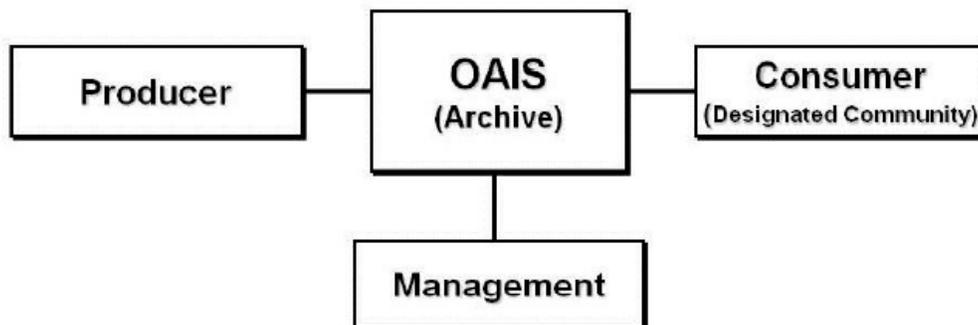
Por último, con este diagrama nivel 3 se tienen conceptualizados el ¿QUÉ? respecto al entorno que envuelve a la edición, por lo que con estos tres niveles de especificación se han determinado los elementos que forman parte del ambiente del modelo y sus relaciones, de ahí que a continuación se hablará del propósito de estos elementos y su influencia en la inserción de metadatos para preservación digital a un nivel abstracto.

### 3.4. ABSTRACCIÓN FRENTE A DETALLE.

Como ya se mencionó, en esta sección se hablará de cómo participan en la preservación digital los elementos que hasta ahora se han identificado como parte del contexto de la edición o ciclo de vida del libro, de tal forma que sea posible establecer, sin entrar en detalles, los factores determinantes que ayudan al resguardo a largo plazo de la publicación.

Para ubicar la participación o relación entre las clases identificadas en el modelo y la preservación digital, es conveniente recordar el ambiente establecido dentro del modelo OAIS, en el cual, tenemos tanto al *productor* como creador del objeto, al *consumidor* o comunidad designada, al *gestor* o responsable del resguardo a largo plazo y al archivo o espacio en el que se lleva a cabo el proceso de preservación, ver Figura 40 (OCLC & Lavoie, 2014, p. 9).

Figura 40. Ambiente OAIS.



Fuente: OCLC & Lavoie, 2014, p. 9.

Por otro lado, en el modelo que aquí se presenta, se tienen identificados varios participantes en el proceso editorial, dentro de los cuales se encuentra el *Ente-Editor*, *Ente-Autor*, además del *Ente-Archivo* y *Ente-Consumidor*, por lo que ahora conviene mirarlos desde la perspectiva que OAIS plantea y para eso es necesario establecer los alcances de estos participantes tanto en la edición como en el ámbito de preservación digital.

Por lo anterior, es claro que dentro del proceso editorial los involucrados tienen dos formas de participación, aquellos que participan de forma **directa** pues llevan a cabo la producción como tal y los que participan de forma **indirecta**, pues no ejecutan el proceso, sólo son el referente pues delimitan o impactan en los criterios, políticas, estándares, tendencias o necesidades establecidas para la publicación, de ahí que en el diagrama nivel 4 del modelo, se presentan dos paquetes de clases identificadas como, el paquete *Ente-Productor* y paquete *Ente-Regulador*, ver Figura 41.

Con este diagrama UML de empaquetamiento se puede decir que los elementos del ambiente de preservación digital también se ven representados, gracias a las relaciones entre estos dos grandes paquetes del nivel 4 del del modelo y la publicación:

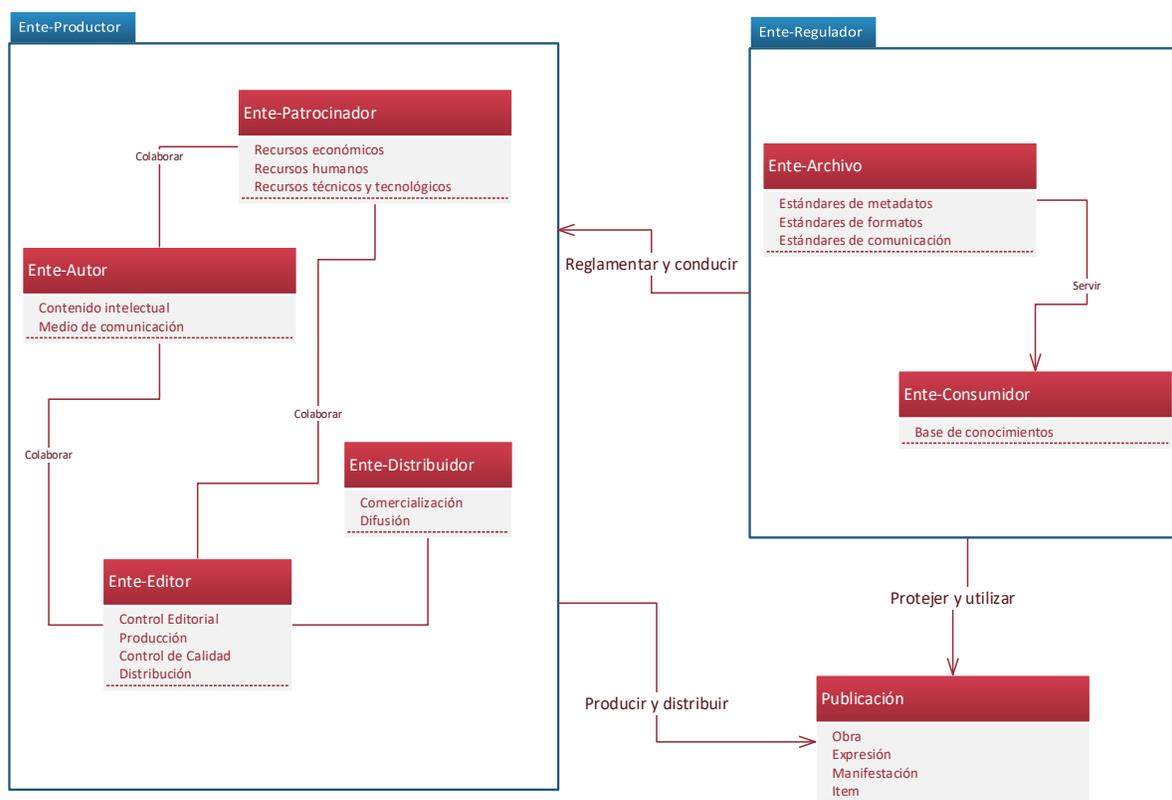
- El paquete *Ente-Regulador* va a *reglamentar y conducir* al paquete *Ente-Productor*
- El paquete *Ente-Productor* va a *producir y distribuir* a la *Publicación*
- El paquete *Ente-Regulador* va a *proteger y utilizar* a la *publicación*

En donde el paquete *Ente-Productor* incluye a cuatro clases: *Ente-Patrocinador*, *Ente-Distribuidor*, *Ente-Autor* y *Ente-Editor*. Es muy importante aclarar que se trata de un planteamiento conceptual y por lo tanto sólo se abstraen para identificar la diversidad de papeles que intervienen, lo cual significa que, en la vida real, una misma persona u organización puede fungir como el paquete de clases *Ente-Productor* y *financiar*, crea su contenido, expresarlo, materializarlo y publicarlo. Lo importante es entender las responsabilidades y compromisos distintos, por lo que, en caso de ser representados por

una misma persona u organización, lo esencial es tomar en cuenta todas las perspectivas funcionales implicadas en la producción, con lo cual se amplía la especificación del modelo OAIS pues el modelo aquí desarrollado nos permite ver todo lo que hay detrás del concepto de *productor*.

Por su parte el *Ente-Regulador* empaqueta al *Ente-Archivo* y *Ente-Consumidor*, cabe aclarar que esto obedece también a lo planteado en el modelo OAIS, pues este se establece la relevancia que tiene el depósito de los productores con características que ayuden a la preservación, tanto como sea posible, por lo que una de las encomiendas dentro de la especificación de OAIS, es establecer las recepciones basadas en acuerdos entre el *Productor* y el *gestor del archivo*, de tal forma que estas contemplen metadatos y objetos estandarizados. También el mismo modelo OAIS, habla de la importancia de que el *Archivo* considere las necesidades de su comunidad designada al momento de establecer los alcances en la preservación.

Figura 41. Modelo, nivel 4, paquetes de clases para la identificación de Ente productor y Ente regulador.



Fuente: La autora diciembre 2019.

Además, con lo anterior se refuerzan las conclusiones identificadas al aplicar las recomendaciones de Ingeniería de Software en el proceso editorial propuesto en este trabajo de investigación, que se refieren a la importancia de considerar dentro de los alcances de la publicación las regulaciones establecidas por instituciones encargadas del

resguardo y las necesidades del lector, de ahí que, en el nivel 4 del modelo se contempla una relación con el *Ente-Productor* de reglamentación pues es necesario que los acuerdos vayan más allá de una aplicación por voluntad, con lo cual se pueda garantizar la aplicación de estándares mínimos definidos por el *Ente-Archivo*.

Por otro lado, la relación de conducción entre el *Ente-Productor* y *Ente-Regulador* se enfoca en el hecho de que la publicación se dirija a satisfacer las nuevas necesidades de lectura y servicio, con lo cual se habla de que la producción de publicaciones digitales puede conducirse por la información que el *Ente-Regulador* genera como *Ente-Archivo* respecto a su comunidad designada o *Ente-Consumidor*.

Así pues, el paquete de la clase Ente-Regulador crea una simbiosis entre el *ente-Archivo* y el *Ente-Consumidor* centrada en el servicio que otorga el primero al segundo, con lo cual es posible crear bases de conocimiento para definir normativas, factores legales y estándares para fines de preservación a largo plazo, que le permitan proteger y utilizar a la publicación, tal como se recomienda en OAIS.

El siguiente paso una vez identificado el alcance de los participantes tanto en el proceso editorial como en la preservación digital, es identificar cómo debe ser la composición de una publicación preparada para preservación a largo plazo y para eso es recomendable recordar los paquetes de información que maneja el modelo OAIS, ver Figura 7, del capítulo I.

De ahí que, se tiene especificado dentro de OAIS que el *productor* es quién deposita en el sistema de archivo, lo que en su momento va a transformarse en un paquete de información SIP<sup>172</sup>, el cual es construido o armado dentro del sistema, gracias a la información que el mismo productor proporciona, para convertirse en el paquete AIP<sup>173</sup>, que en esencia es el objeto preservado a largo plazo, además, el AIP también permite la creación de paquetes DIP<sup>174</sup>, los cuales son utilizados para proporcionar el acceso a los diferentes consumidores.

Esto nos lleva a analizar lo que OAIS señala respecto al producto generado por el *Ente-Productor*, conocido como paquete de información, el cual en estricto sentido debe contar con lo necesario para conformarse como un paquete SIP y después, como un paquete AIP, que será almacenado en un ambiente de preservación, pues es la información que el *Ente-Archivo* recibe, organiza y proteger el mayor tiempo posible, ver Figura 42.

En el modelo de información OAIS, el AIP está conformado por un *objeto dato contenido*<sup>175</sup> (CDO), el cual, en esencia es el objeto digital, además de la *información de representación*<sup>176</sup>

---

<sup>172</sup> Submission Information Package

<sup>173</sup> Archival Information Package

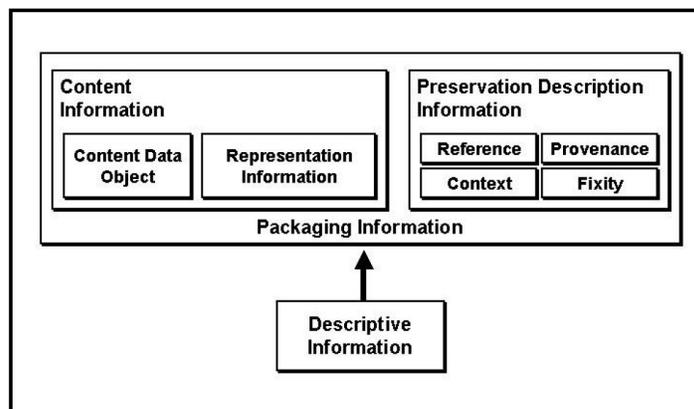
<sup>174</sup> Dissemination Information Package

<sup>175</sup> Content Data Object

<sup>176</sup> Representation Information

(RI) que es la que nos permite entender qué es y cómo se constituye el *CDO*, ambos idealmente deben ser proporcionados por el *Ente-Productor*, para constituir la *información de contenido*<sup>177</sup> (CI). Además, cada AIP contiene *Información descriptiva de preservación*, la cual se refiere a los procesos aplicados al CI, concentradas estos dos como *Información de empaquetado* y su correspondiente *Información Descriptiva* del propio AIP (Gartner & Lavoie, 2013).

Figura 42. (AIP) Paquete de Información de Archivo OAIS.



Fuente: Gartner & Lavoie, 2013.

Retomemos entonces el diagrama del nivel 1 del modelo que habla de los componentes de la *publicación*, que para efectos de esta investigación podemos identificar como la representación del CDO<sup>178</sup> en OAIS, pues simboliza al elemento que requerimos preservar, que en esencia encapsula cuatro componentes: *Obra, expresión, manifestación e ítem*. En donde además cada uno de estos, se caracteriza por generar información que ayuda a entender la forma en que se compone la publicación, de ahí que estas características puedan compararse con la *Información de representación*<sup>179</sup> (RI) que OAIS señala como aquello que permite entender al objeto y que en esencia incluye datos derivados del *Ente-Productor*.

Todo esto, se ve representado en el diagrama nivel 5 del modelo que en esencia resalta el hecho de que la *Información de representación* debe estar incorporada como *metadatos y documentación* derivada de los diversos componentes de la *publicación*, por lo que las características de cada componente nos permiten identificar el momento en el que es posible generar la *Información de Representación* o RI, que al final está contenida en el ítem, lo cual concuerda con el ideal en el modelo OAIS, en el entendido de que el *Ente-Productor* entregue la información suficiente para la conformación del paquete SIP, ver Figura 43.

<sup>177</sup> Content Information

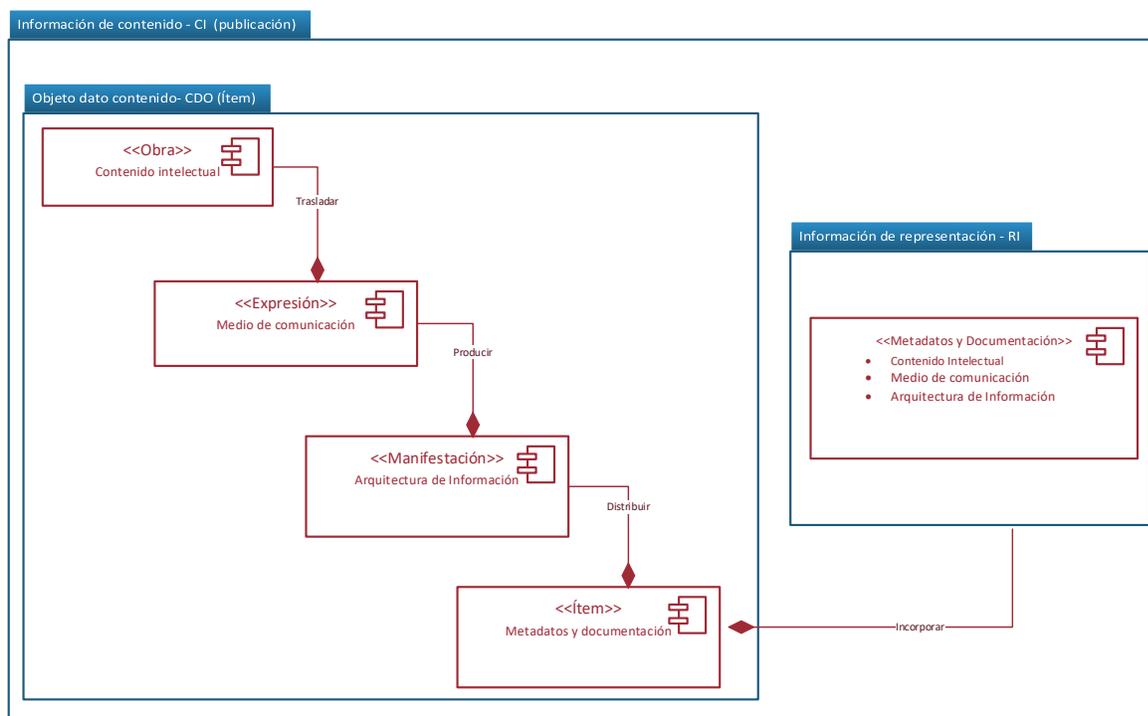
<sup>178</sup> Content Data Object

<sup>179</sup> Representation Information

Sin embargo, para que el ítem de la publicación, cumpla con los elementos necesarios para preservación, es necesario recordar lo señalado al final del capítulo anterior, respecto a lo que permite mejorar la calidad de la publicación digital y en consecuencia preparar un producto editorial para preservación:

- a. **Establecer alcance y composición de la publicación.** Que incluye la especificación de la estructura de la publicación y la Identificación de recursos a utilizar de acuerdo con el alcance deseado.
- b. **Definir criterios básicos de calidad del software.** Que contempla el uso de estándares y la reutilización recursos o prácticas generadas previamente en la organización o por la comunidad, de acuerdo con el alcance deseado.
- c. **Trabajo colaborativo de todos los involucrados.** Que contempla solicitar o incluir el apoyo de expertos para cada actividad nueva y desde luego contempla los pasos ya conocidos para los procesos editoriales tradicionales, considerando lo necesario a pesar de que se trate de publicaciones impresas que se publican también en formato digital.
- d. **Verificación y control del proceso de producción y de la publicación.** Que incluye la documentación del proceso y el seguimiento para garantizar la calidad esperada en el resultado.

Figura 43. Modelo, nivel 5, componentes empaquetados del Contenido Informativo (CI) de la publicación.



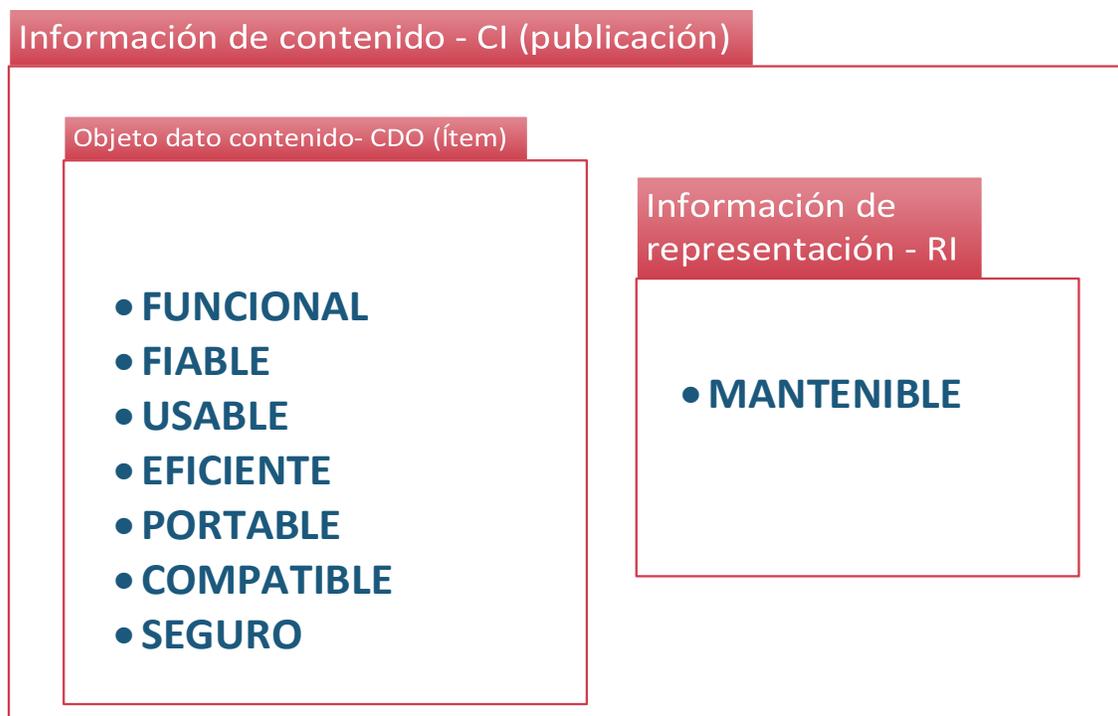
Fuente: La autora diciembre 2019.

De lo anterior vale recalcar, que con el avance que hasta ahora se tiene en la especificación del modelo objeto de esta investigación, se ha cumplido con el inciso a) *Establecer alcance y composición de la publicación*, pues los diagramas planteados hasta ahora, identifican los elementos del entorno y la composición de la publicación, en dónde no se especifica cuál es el contenido exacto, sólo se delimitan los componentes abstractos y generales que conforman al libro digital y que desde esa perspectiva, permitirían integrar las características que ayuden a su preservación.

Para el siguiente punto, b) *Definir criterios básicos de calidad del software*, retomemos los atributos de calidad mencionados en el capítulo anterior: **Funcionalidad, Fiabilidad, Usabilidad, Eficiencia, Mantenimiento, Portabilidad, Compatibilidad y Seguridad**. Para ubicarlos dentro de la *Información de Contenido*, como se muestra en el diagrama nivel 6 del modelo, ver Figura 44.

Para esto, es importante recordar que la mayoría de los atributos de calidad de software califican al producto final en su comportamiento individual y al interactuar con otro software, aplicación y/o usuario final y en es este caso, el producto final es el ítem de la publicación, de ahí que el Objeto dato contenido o CDO concentra la mayoría de estos atributos que ayudan a describirlo como objeto digital.

Figura 44. Modelo, nivel 6, diagrama de paquetes que muestra los atributos de un objeto digital.



Fuente: La autora diciembre 2019.

Sin embargo, existe un atributo conocido como *mantenimiento*, que se refiere a la información y capacidad que el software debe cumplir para que sus usuarios puedan llevar

a cabo mejoras y actualizaciones que permitan alargar la vida útil del mismo, de ahí que, el *mantenimiento* en términos de Ingeniería de Software se concentra en la vida rentable del producto ya que en términos operativos, este se mantiene siempre y cuando su funcionalidad cumpla con la necesidades de la organización, sin embargo en términos de preservación digital, el concepto de mantenimiento podría ser corto en su alcance pues la preservación del libro digital, coloca el contenido de la obra por encima de la funcionalidad, de ahí que, de no ser posible el rescate o conservación de la funcionalidad, si es indispensable la preservación del contenido.

Por lo tanto, el atributo mantenimiento desde la perspectiva del libro digital requiere contemplar elementos que prioricen la preservación del contenido incorporando mecanismos que faciliten el traslado a otros formatos o medios y esto también se relaciona con el atributo de portabilidad, para prever esta situación.

De lo anterior que, la perspectiva del mantenimiento es lo que mejor se aproxima a la sobrevivencia del producto digital, por lo que para efectos de este trabajo de investigación se considerará como punto de partida, ampliando el alcance del mantenimiento a una supervivencia mayor, basado en que los principios que rigen el mantenimiento también van a beneficiar a la preservación digital, pues principalmente se centra en generar la documentación suficiente y adecuada para el entendimiento y manejo del software una vez que se encuentra en uso.

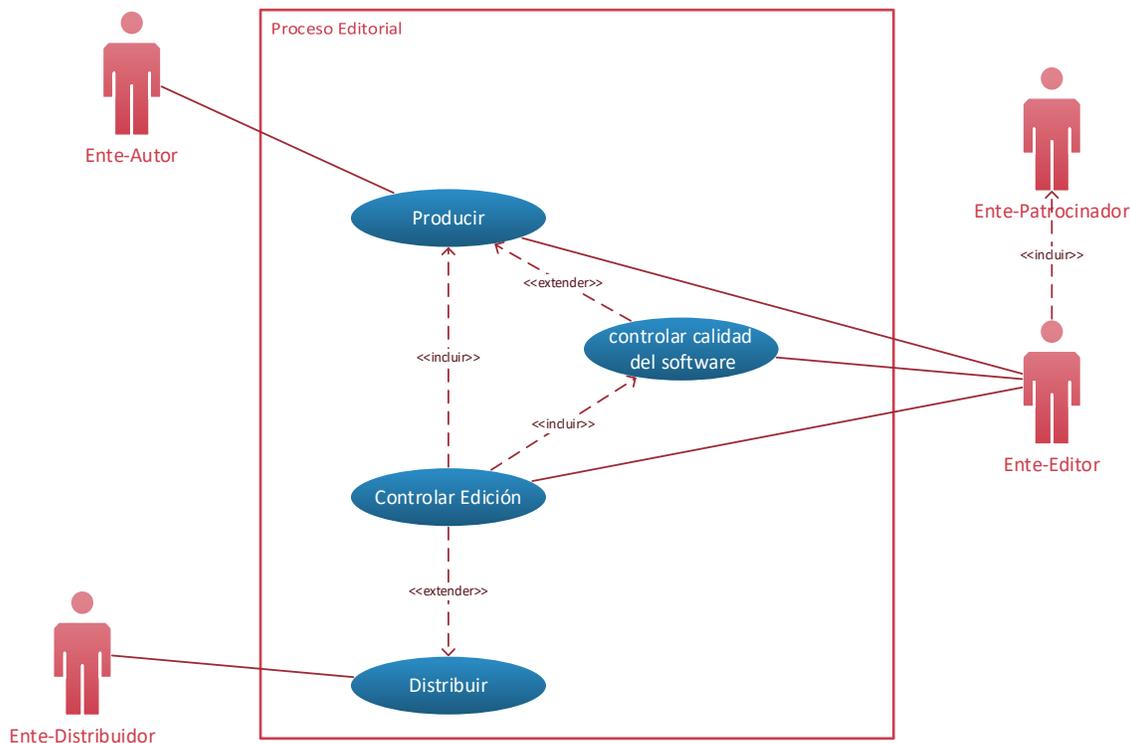
A partir de todo esto, entonces, se recomienda que durante el proceso de edición es necesario incorporar elementos de verificación que permitan garantizar que la publicación cumpla con los siguientes criterios en su construcción, para asegurar la integración de metadatos técnicos referentes a:

- La funcionalidad, que debe estar descrita y documentada desde que se construye.
- Además, la construcción del libro digital debe considera el uso de herramientas estándares para garantizar su fiabilidad, usabilidad, eficiencia, portabilidad, compatibilidad y seguridad.
- La portabilidad debe ser parte del alcance de la publicación, pues ayuda a garantizar al mantenimiento de la publicación.
- Y es necesario acompañar al objeto, con *información de representación*, de tal forma que esta ayude en su mantenimiento a largo plazo.

El siguiente elemento que procura la preservación de la publicación digital desde la perspectiva de la Ingeniería de Software, es el que se refiere al inciso *d) Verificación y control del proceso de producción y de la publicación*, del listado antes citado, que se centra en el seguimiento y cuidado del proceso editorial, para lo cual se retoma es conveniente retomar el proceso generado en el capítulo anterior y representarlo de forma aún más abstracta, utilizando un diagrama UML de casos de uso, en donde cada etapa del proceso

se abstrae como un caso de uso y las clases pasan a ser actores en cada caso de uso, ver Figura 45, diagrama nivel 7 del modelo.

Figura 45. Modelo, nivel 7, diagrama de casos de uso para representar el proceso editorial y sus actores.



Fuente: La autora diciembre 2019.

De esta forma es posible visualizar de forma muy general la asociación que existe entre el *Ente-Editor* y resto de los participantes y las cuatro etapas definidas del proceso editorial propuesto en este trabajo de investigación, que además contiene dos tipos de control, el de la edición en general, que como ya se vio se refiere al proceso en sí y la calidad del software, lo cual además permite ubicar que la revisión y verificación de los atributos de calidad de la publicación se concentra en la etapa de *calidad del software* que se extienden a la etapa de *producción*, por lo que éste es el lugar en el que se va a generar la *Información de Contenido* referida en OAIS, que contempla tanto a los metadatos descriptivos como a los técnicos o atributos de calidad de software.

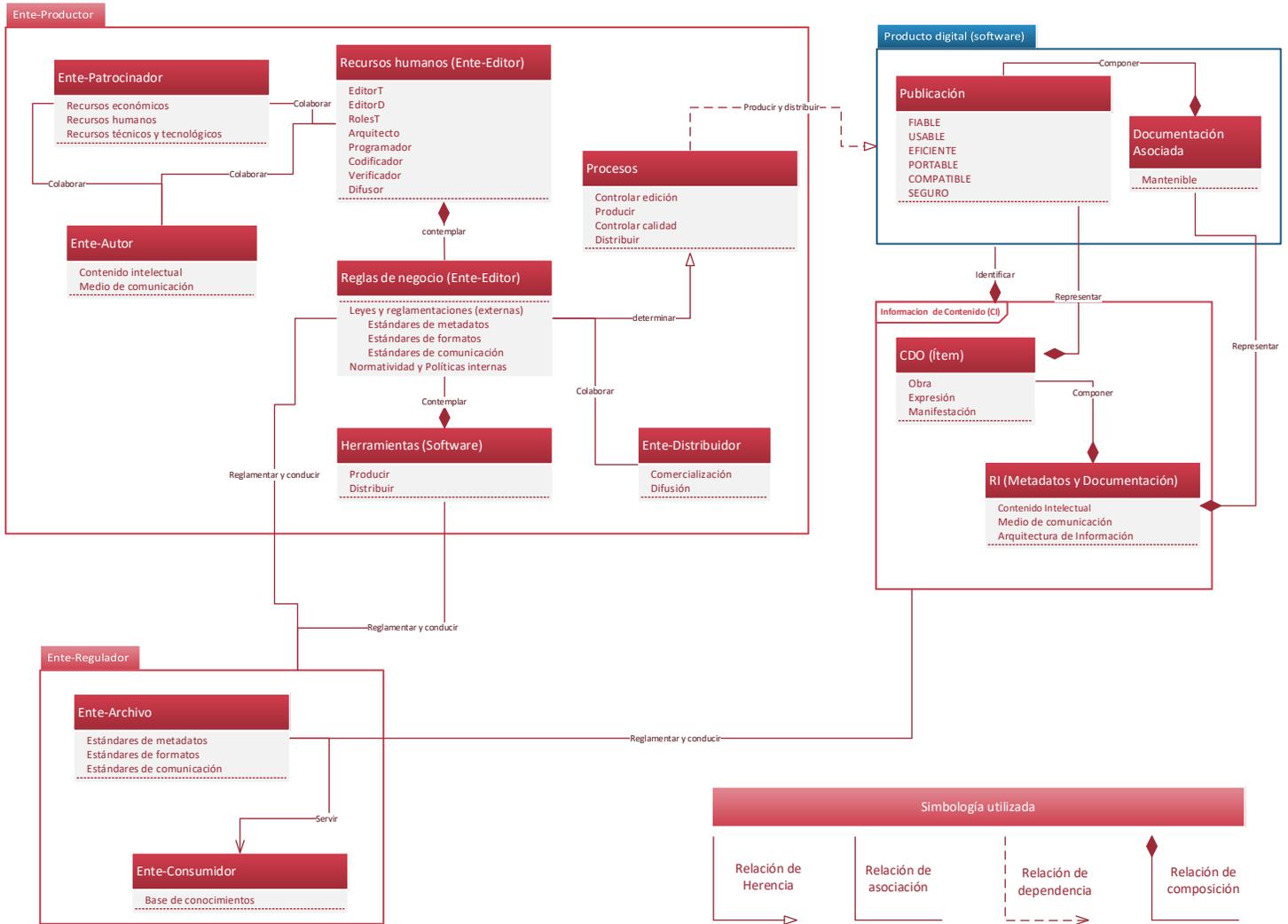
El último elemento relacionado con la preservación de la publicación se señala en el inciso *c) Trabajo colaborativo de todos los involucrados*. Para esto es necesario recordar los factores que se identificaron en el capítulo anterior como participantes desde el punto de vista organizacional, en el ciclo de vida del libro y que envuelven la edición:

- a. **El software.**
- b. **Los procedimientos.**

- c. Las reglas de negocio.
- d. Los recursos humanos.
- e. La documentación asociada.

Para llevar estos elementos, al diagrama nivel 8 del modelo, cada uno de los factores se va a caracterizar en un diagrama de clases, relacionado con los conceptos vistos a lo largo del modelo y en el capítulo anterior, con el proceso de edición propuesto, ver Figura 46 .

Figura 46. Modelo, nivel 8, diagrama de paquetes para representar los factores que intervienen en el proceso editorial.



Fuente: La autora diciembre 2019.

En este diagrama el *software* tiene dos representaciones, primero como el libro digital que debe cumplir con ciertos atributos de calidad y está compuesto como un paquete de información y, en segundo lugar, el software se refiere a las herramientas que componen las reglas del negocio que sirven para la producción y distribución de los libros digitales,

que, de igual forma, debe apegarse a ciertos estándares para contribuir a la preservación de la publicación.

Los *procedimientos*, son lo que se establecen en la edición y que agrupa acciones como: *controlar edición, producir, controlar calidad del software y distribuir*. Estos además están determinados por *las reglas de negocio* y encargados de producir y distribuir el libro digital o publicación.

Las *reglas de negocio* representan la esencia del Ente-Editor, pues contienen las leyes y reglamentaciones externas a las que se apega y que están conducidas por una *Entidad-Reguladora*. Además, contienen las normas y políticas internas, de ahí que se componen tanto de los recursos humanos como de las herramientas utilizadas y se generan a partir de la colaboración del *Ente-Patrocinador* y el *Ente-Distribuidor*.

Los recursos humanos integran a todos los participantes de los procesos de ahí que forman parte de las reglas de negociación y además pueden integrar a colaboradores financiados por el *Ente-Patrocinador* y *Ente-Autor*.

Por último, la *documentación asociada* se genera durante la producción en el proceso de edición y se refiere a los metadatos para preservación que incluye tanto metadatos descriptivos como técnicos y administrativos, que forman parte del producto digital que se distribuye, pues como ya se vio dentro del paquete de información OAI, va a contener la información de representación o RI, requerida para entender al objeto digital, que en este caso se trata de un libro y que como ya se vio, representa la capacidad de que el producto sea mantenible.

Este es el último nivel del modelo agrupa todos los elementos analizados desde el contexto, hasta la concepción de lo que es el libro digital y los elementos que ayudan a evaluar su calidad, como software y como paquete de información de preservación, por lo que en el siguiente capítulo se analizará la forma en que este modelo puede ser aplicado o implementado ya sea por una entidad editora o servir de referencia para que una entidad de resguardo defina lineamientos para los editores o productores de libros digitales.

### 3.5. RECOMENDACIONES DE APLICACIÓN E IMPLEMENTACIÓN DEL MODELO.

En esta sección se explicarán las recomendaciones de aplicación del modelo desde dos puntos de vista, desde el punto de vista del editor o productora de libros digitales y desde la mirada de una entidad receptora del Depósito Legal encargada del resguardo y preservación del libro digital, sin embargo, antes es necesario especificar cuáles son los elementos que caracterizan al Modelo de integración de metadatos para preservación en la edición del libro digital desarrollado aquí:

1. Establece un conjunto de supuestos respecto a la edición del libro digital que pueden extenderse a otro tipo de publicaciones.

2. Describe el sistema que gira alrededor de la edición del libro digital, de tal forma que plantea la estructura general del proceso y el sistema organizacional que lo respalda además de exhibir las propiedades básicas del libro digital.
3. Sirve como una aproximación de lo que implica la edición del libro digital.
4. Y establece aspectos para la calidad del libro digital desde una perspectiva del software que abarcan su estructura y composición digital, de ahí que es posible identificar una clasificación de los metadatos técnicos derivados de su esencia como software.

---

### 3.5.1. DE LA APLICACIÓN COMO EDITOR

De lo anterior que, el modelo aquí expresado le permite al editor:

1. Identificar rasgos o características generales del libro digital como software
2. Identificar los procesos generales que forman parte de la edición del libro digital
3. Identificar la injerencia e importancia de entidades regulatorias en la edición
4. Identifica cuáles son los metadatos técnicos con base en los atributos de calidad de software que debe considerar para la preservación de su obra.
5. Muestra los momentos en los que durante la edición se definen los alcances de los metadatos y se generan, como documentación asociada a la publicación.
6. Presenta un proceso de edición para el libro digital que puede ser adaptable a las necesidades del editor.
7. Presenta al libro como software y la esencia de la Ingeniería de Software por lo que es posible que el editor mire al libro desde otra perspectiva
8. Resalta la importancia de las reglas del negocio en la edición del libro digital
9. Resalta la necesidad de especificar procedimientos, pues estos determinan la calidad del libro desde su mirada como software
10. Introduce la importancia del concepto de mantenimiento como parte de la preservación digital del libro.

Lo que el modelo NO establece:

1. Pasos concretos u obligatorios
2. Metadatos concretos u obligatorios

Por lo tanto, los editores pueden ocupar el modelo que aquí se presenta para dirigir su mirada al interior de sus negocios y reflexionar acerca de las *reglas de negocio* con las cuales trabajan para identificar, sí sus procesos y la normatividad interna contempla al menos las siguientes recomendaciones:

#### 1. En cuanto al control de la edición.

- a. Identificar que se tengan los recursos suficientes: humanos, económicos y tecnológicos para llevar a cabo la producción del libro digital.

- b. Identificar que los recursos humanos cuenten con las habilidades necesarias considerando el proceso editorial de libros digitales como un proceso de producción de software.
  - c. Considerar la participación definitiva de un editor digital, que apoye en control de la calidad del libro como producto digital.
  - d. Considerar el contacto con entidades encargadas de la preservación digital de libros en su localidad, para atender a las necesidades establecidas por las mismas en pro de la permanencia sus publicaciones.
  - e. Considerar durante la edición, las necesidades de sus lectores, distribuidores y entidades encargadas del resguardo de las publicaciones digitales.
- 2. En cuanto proceso.**
- a. Establecer un proceso editorial que además de atender la calidad del contenido intelectual ponga atención en la calidad del libro como software, considerando los atributos referidos en el modelo.
  - b. Tener por escrito las características que precisan el alcance de la publicación, para tratar de confirmar que las necesidades reales son igual a las expresadas y especificadas, de tal forma que al final, estas mismas sean necesidades cumplidas.
- 3. En cuanto al libro digital.**
- a. El alcance del libro digital debe considerar siempre documentación o metadatos que hablen de su contenido intelectual, medio de comunicación y arquitectura de información, que en esencia refieren el alcance completo de la publicación.
  - b. Los criterios de calidad establecidos para el libro digital deben contemplar metadatos técnicos que describan la funcionalidad, compatibilidad, seguridad, portabilidad, eficiencia, usabilidad y mantenimiento, además de contemplar los elementos que ayuden al cumplimiento de estos criterios de calidad.
  - c. El libro digital no sólo debe ser atendido en su producción, además ahora requiere contemplar servicios que ayuden durante la distribución y difusión con el soporte adecuado a sus lectores y otros, interesados en la publicación.

---

### 3.5.2. DE LA APLICACIÓN EN ENTIDADES QUE RESGUARDAN LIBROS DIGITALES

Respecto a la utilidad que tiene el modelo para las instituciones encargadas del resguardo de libros digitales, está en primera instancia el establecimiento de una relación de participación en el proceso editorial como actor indirecto, en la definición del alcance de las publicaciones como una entidad que promueve la permanencia de la publicación a largo plazo.

También se tiene identificado dentro del modelo la composición del libro digital desde una perspectiva de preservación digital, por lo cual sirve a instituciones como la Biblioteca Nacional de México, de punto de referencia, para establecer las características generales que las entregas de los productores debieran contemplar.

En este sentido, el modelo sirve a las entidades responsables del resguardo dado que presenta los siguientes elementos:

- a. El libro digital se compone como software, de ahí que la especificación de su arquitectura es parte importante para la preservación a largo plazo.
- e. Con base en los atributos de calidad de software del libro digital, es recomendable que la evaluación en cuanto a contenido de metadatos para preservación considere los siguientes elementos: Funcionalidad, Fiabilidad, Mantenimiento, Portabilidad, Compatibilidad y Seguridad.
- h. La funcionalidad se refiere a la interacción que la publicación contempla con sus usuarios o lectores, por lo que es recomendable que la documentación de la funcionalidad especificada en metadatos permita su interacción en ambientes de preservación digital.
- i. La fiabilidad, compatibilidad, portabilidad y mantenimiento son atributos estrechamente relacionados con la sobrevivencia del libro digital en un ambiente de preservación.
- j. La seguridad contempla las estrategias empleadas durante la construcción y aquellas implementadas para el uso del libro digital, por lo tanto, es recomendable que el material contenga los metadatos necesarios para controlar las restricciones y cuidar que se cumpla con los derechos otorgados a los lectores o usuarios de la obra.
- e. Es necesario contar con medios que apoyen o soporten el uso del libro digital, desde manuales de usuario hasta servicios en línea que orienten o resuelvan inquietudes de los lectores, sin embargo, la integración de esta información en metadatos también será fundamental.

Por lo tanto, la perspectiva que otorga el modelo a las instituciones de resguardo es meramente referencial, sin embargo, ayuda a visualizar los nuevos elementos que forman parte de la composición del libro digital y que, por su naturaleza desde la perspectiva del software, le exigen criterios de calidad que van más allá de su contenido intelectual, por lo que evidencia la importancia y alcance de los metadatos técnicos que ayudarán al cuidado y protección de la publicación.

## CONCLUSIONES.

El objetivo de esta investigación fue, generar un modelo de integración de metadatos para preservación en la edición del libro digital, sin embargo, antes del desarrollo del modelo fue necesario entender varios conceptos, entre ellos el del libro digital, por lo cual, en el primer capítulo, se identificó que éste tiene una función como emisor de ideas tanto en su contenido como en su forma iconotextual, derivado de su función como objeto libro.

Además, para el análisis de su naturaleza digital se revisó cronológicamente su origen como producto del avance en las tecnologías de la información y comunicación, con lo cual se identificaron tres perspectivas que han marcado su presencia intangible, primero como texto digital en donde gracias a las computadoras, internet y la web, es posible el intercambio de las palabras e ideas contenidas en el libro, incluso desde antes de ser publicado pues ahora el texto nacen digital incluso para un libro impreso.

La segunda perspectiva de origen es la aparición de formatos de archivo y aplicaciones de software que poco a poco, han permitido la configuración iconotextual del libro en el ámbito virtual y que son, el punto de partida para otorgan funcionalidades y/o nuevas dinámicas de lectura a los usuarios. Y el tercer punto, es el desarrollo de dispositivos de lectura, pues estos desde sus inicios buscaron completar la metáfora del libro, materializando su esencia intangible y ampliando las posibilidades para contener más de un libro en el mismo dispositivo, lo cual también ha generado bibliotecas portátiles.

Dicha revisión cronológica junto con el análisis de conceptos y perspectivas descriptivas, permitieron demostrar la naturaleza del libro como objeto digital y definirlo para efectos de esta investigación como:

*software, objeto digital o tipo especial de información constituido en uno o más formatos de archivo de computadora con comportamiento computacional para abstraer la idea del libro en el mundo material, que requieren de un sistema informático (con nuevas dinámicas en su ciclo de vida y cadena de valor) para lograr interactuar con el mundo físico, es decir, dispositivos de hardware, otras piezas de software (programas de computadora y otros objetos digitales) y documentación asociada como metadatos.*

Esta definición también permitió establecer el primer punto de partida del porqué la Ingeniería de Software puede tener influencia en la edición del libro digital, ya que, al hablar de un sistema informático, las dinámicas de su ciclo de vida y cadena de valor se ven afectadas, primero por el alcance que puede tener el libro en cuanto a funcionalidades y luego por la intervención de nuevos perfiles profesionales, lo cual está relacionado con la segunda hipótesis de esta investigación, sin embargo, es un tema que se explica más adelante.

Otro concepto desarrollado al final del primer capítulo, fue lo referente a los elementos que permiten la preservación, de ahí que, se revisaron los cambios generados a nivel mundial en la legislación del Depósito Legal, ya que este es el principal mecanismo de recolección de libros para conservar en entidades depositarias como bibliotecas y archivos, con lo cual se identificó que muchos países han generado importantes cambios en dicha legislación para considerar la preservación digital con la participación de los editores y entidades depositarias. Esto, además, permitió confirmar parte de la primera hipótesis de esta investigación, ya que los casos revisados aseguraron que en la práctica es posible que los editores entreguen metadatos junto con sus publicaciones.

También se revisaron los avances en México respecto al Depósito Legal y la preservación digital, con lo cual se identificó que aún hay mucho trabajo por hacer, pues a pesar de que ya se tienen algunos cambios en la legislación, la preservación digital aún no figura como un concepto y tampoco existe un trabajo conjunto con los editores en este tema, sin embargo, es un asunto que sigue tratándose en la Biblioteca Nacional de México, por lo que es posible que las cosas sigan cambiando y mejorando.

Además, se profundizó en la preservación digital, para entender cuál es su alcance y conocer qué organizaciones a nivel mundial están trabajando en el tema y qué es lo que han avanzado en el desarrollo de estrategias para la conservación de libros digitales, de lo cual se identificó la generación de metadatos, como uno de los puntos sustanciales.

Por otro lado, al revisar con detalle el modelo OAIS, un estándar principalmente para documentos de archivo pero aplicable a la preservación de libros digitales, se identificó que también este modelo resalta la importancia de los productores de objetos digitales, que en este caso son los editores y además, explica con detalle cuáles son los **metadatos de preservación** generados durante el proceso y los **metadatos para preservación** que incluyen otros subconjuntos de metadatos, como los descriptivos, técnicos y legales entre otros.

Todo lo anterior, permitió identificar la importancia que tienen en la preservación digital la participación de los editores, las entidades depositarias y los metadatos pues tanto en el modelo OAIS como en las prácticas expuestas respecto al Depósito Legal en otros países, dejaron evidencia para confirmar esto, que es parte, de la primera hipótesis de esta investigación. Con esto, también se dio respuesta a la primera pregunta de investigación, ya que el editor como actor principal en la cadena de valor del libro, es quien tiene la posibilidad de generar los metadatos para preservación, de ahí que, países como Alemania muy adelantado en la preservación del libro digital, contemplan la entrega de metadatos directamente de los editores y las entidades receptoras del depósito legal, regulan todo el proceso.

Además, como ya se mencionó, se identificó que con la evolución del libro en el ámbito digital se han transformado diversas dinámicas alrededor de él, entre ellas su ciclo de vida

y su cadena de valor, dentro de las cuales se encuentra la edición, pues, por ejemplo, ahora los autores, pueden prescindir de los editores y distribuidores, creando sus propias publicaciones y exponiéndolas en la web para la lectura.

De ahí que, el segundo capítulo de esta investigación se concentró en el análisis de la edición como parte del ciclo de vida del libro y su cadena de valor, primero para entender cuáles son las nuevas dinámicas identificadas, que involucran prácticas de Ingeniería de Software que como se dijo antes se derivan de la naturaleza del libro como software, por lo cual los editores van a requerir de integrar nuevos perfiles profesionales durante la edición y además establecer nuevas dinámicas para la creación y distribución de sus productos, pues incluso es necesario atender necesidades que pueden demandar los lectores con el uso del libro posterior a la compra o adquisición.

Lo anterior deriva en la identificación de una etapa en la edición que en esta investigación se denominó como producción, que concentra la especificación de la arquitectura de información, construcción de funcionalidades y codificación del libro digital, en la cual, una parte del trabajo es la generación de metadatos descriptivos, de ahí que a partir de los conceptos de ingeniería de software, se identificaron atributos de calidad de software que describen características propias de un libro digital y que permiten caracterizar a los metadatos básicos del mismo, durante este proceso. Con lo cual se terminó de confirmar la primera hipótesis de investigación.

Así pues, lo siguiente en esta investigación fue identificar cuáles son los metadatos básicos que se deben generar durante la edición del libro digital y para eso, lo primero fue identificar qué implica el proceso de edición del libro digital a partir del estudio de casos tipo y por oportunidad con lo cual, se diseñó un proceso editorial basado en cuatro ejemplos seleccionados por representar perspectivas de edición diferentes, a los cuales, se añadieron teorías fundamentadas en la Ingeniería de software para generar un proceso de edición que sirvió como base del modelo aquí descrito.

Con esto fue posible dar respuesta a las dos últimas preguntas de investigación pues se identificó que los metadatos básicos que deben generarse durante la edición del libro digital son los que se refieren a su funcionalidad, usabilidad, fiabilidad, eficiencia, portabilidad, compatibilidad, seguridad y mantenibilidad. Por lo que este modelo hace una clasificación de los metadatos a partir de atributos del libro como software.

La última parte de esta investigación se concentró en el diseño del modelo de integración de metadatos para preservación en la edición del libro digital, con lo cual se dio respuesta a la última hipótesis planteada en este trabajo, pues se basó en dos perspectivas documentales importantes FRBR y el modelo OAIS, además de la aplicación de conceptos derivados de la ingeniería de software integrados en el proceso de edición que se diseñó, a partir de la concepción del libro digital como software.

De lo anterior que, con el modelo, se representó el papel regulador que juegan las entidades receptoras del depósito legal, la participación de los editores y los metadatos básicos que pueden ser generados durante el proceso de edición, con la aplicación de buenas prácticas de Ingeniería de Software, lo cual permitió comprobar la dos hipótesis planteadas en este trabajo de investigación.

Sin embargo, cabe mencionar que a pesar de que existen diversos esquemas de metadatos para la descripción de características de contenido y estructurales de los objetos digitales, aún no existe una especificación o un esquemas de metadatos que se adecue a los metadatos básicos señalados en esta investigación para el libro digital, de ahí que este punto puede ser una vertiente adicional de estudio, en el futuro, para el desarrollo de un nuevo esquema o la extensión de alguno ya existente, ya que a simple vista, los metadatos aquí descritos parecerían de naturaleza técnica, sin embargo se requiere una trabajo de análisis adicional.

Otra vertiente que se deriva de esta investigación, es el estudio de otros elementos que impactan en la preservación del libro digital y no se restringen a los metadatos, pues como se comentó en su momento, los modelos de negocio empleados por los editores además provocan la pérdida de la propiedad de uso del libro y esto también es un factor crítico para la preservación, ya que el libro no se recibe o cuenta con restricciones de acceso que evita la adecuada preservación por parte de las entidades receptoras, lo cual es un tema no resuelto y que no se toca en esta investigación pues sale del alcance de la misma, sin embargo puede tratarse a partir de algunas de las propiedades del libro como software, referidas como mantenibilidad, seguridad, portabilidad y fiabilidad.

También cabe mencionar que con este trabajo se demuestra la estrecha relación que existe entre la ingeniería de software, la bibliotecología y la preservación digital, pues cada una de estas desde su perspectiva, complementa la visión con la que se pueden analizar los nuevos tipos documentales, que han surgido con el uso de las tecnologías.

En conclusión, fue posible alcanzar los objetivos planteados para este trabajo de investigación, a pesar de que también se encontraron nuevos desafíos y temas a desarrollar derivados de los resultados encontrados por lo que el modelo que aquí se presenta:

1. Tiene un carácter híbrido por basarse en el desarrollo de hipótesis que pertenecen a la preservación digital, la edición de libros y la Ingeniería de Software.
2. Consiste en un conjunto de supuestos respecto a los libros digitales que parten de su esencia como software y que fueron comprobados en el transcurso de la investigación.
1. Describe una publicación digital atribuyéndole una estructura interior, con la cual se explican sus propiedades desde el punto de vista de su preservación y la relevancia que tienen los metadatos.

2. El modelo es útil para aproximarse a los elementos que permiten la integración de metadatos para preservación en la edición del libro digital.
3. Se formuló, desarrolló y designó a partir de la analogía basada en el análisis de algunos casos de procesos de edición descritos en la comunidad editora a partir de advertir propiedades comunes y principios similares.
4. Es un modelo descrito a partir de un lenguaje de modelado gráfico conocido como UML, ampliamente utilizado en la ingeniería de software con el cual es posible modelar aspectos semánticos y de contexto.
5. El modelo da prioridad a la abstracción más que al detalle por lo que sólo contempla elementos esenciales, en este caso para la preservación digital, omitiendo detalles irrelevantes en cuanto a la implementación de procesos.
6. Finalmente, el modelo también prioriza la especificación frente a la implementación, por lo que lo más importante es definir el qué y no el cómo, de ahí que cada institución u organización editora puede adoptar las acciones que mejor le convengan para integrar los metadatos básicos para preservación, aquí señalados.

## REFERENCIAS.

- Achinstein, P. (1967). *Los modelos teóricos* (M. Sancho, Trad.; 2a ed.). Dirección General de Publicaciones UNAM - Coordinación de Humanidades.
- Amazon.com | History & Facts. (s. f.). Encyclopedia Britannica. Recuperado 6 de abril de 2021, de <https://www.britannica.com/topic/Amazoncom>
- Ardito, S. (2000). Electronic Books: To «E» or not to «E»; that is the question. *Searcher: The Magazine for Database Professionals*, 8(1), 28. <https://www.infotoday.com/searcher/apr00/ardito.htm>
- ASALE, R.-, & RAE. (2020a). *Libro | Diccionario de la lengua española*. «Diccionario de la lengua española» - Edición del Tricentenario. <https://dle.rae.es/libro>
- ASALE, & RAE. (2020b). *Libro | Diccionario de la lengua española*. En «*Diccionario de la lengua española*»—Edición del Tricentenario. <https://dle.rae.es/libro>
- Barbier, F. (2015). *Historia del libro* (P. Quesada, Trad.). Alianza Editorial.
- Bhaskar, M. (2014). *La máquina de contenido. Hacia una teoría de la edición desde la imprenta hasta la red digital* (R. M. Rubio, Trad.; Primera edición). Fondo de Cultura Económica, S.A. de C.V. <https://e-books.fondodeculturaeconomica.com/a5be3005-5849-4424-9489-30a094b77c8a>
- Bhaskar, M., & Rubio Ruiz, R. M. (2017). *La máquina de contenido: Hacia una teoría de la edición desde la imprenta hasta la red digital*.
- BIBFRAME-LC. (2016). *Overview of the BIBFRAME 2.0 Model (BIBFRAME - Bibliographic Framework Initiative, Library of Congress)* [Webpage]. <https://www.loc.gov/bibframe/docs/bibframe2-model.html>
- BNM, B. N. de M. (2019). *Tercer informe de actividades 2018-2019* (Anual N.º 3er; p. 153). Instituto de Investigaciones Bibliográficas Biblioteca y Hemeroteca Nacionales de México. [http://www.bnm.unam.mx/files/quienes-somos/informes-anales/tercer\\_informe\\_iib\\_2018.pdf](http://www.bnm.unam.mx/files/quienes-somos/informes-anales/tercer_informe_iib_2018.pdf)
- BNM, B. N. de M. (2021). *Lineamientos de Depósito Legal*. Biblioteca Nacional de México. <https://bnm.iib.unam.mx/index.php/quienes-somos/deposito-legal/lineamientos>
- Boehm, B. W., Brown, J. R., & Lipow, M. (1976). QUANTITATIVE EVALUATION OF SOFTWARE QUALITY. *ICSE '76: Proceedings of the 2nd International Conference on Software Engineering*, 14. <https://dl.acm.org/doi/10.5555/800253.807736>
- Bourque, P., Fairley, R. E., & IEEE Computer Society. (2014). *Guide to the software engineering body of knowledge*. IEEE Comput. Soc.
- Bueno, S. S. (2005). «En un lugar de la “Mancha”...» Procesos de control de calidad del texto, libros de estilo y políticas editoriales. *panace@*, VI(21-22), 16. <http://search.ebscohost.com/login.aspx?direct=true&db=edsoai&AN=edsoai.ocn711547280&lang=es&site=eds-live>
- Burnard, L. (2001). *TEI U5: Etiquetado para el intercambio: Una introducción al TEI* (Q. M. Sánchez, Trad.). Biblioteca Virtual Miguel de Cervantes, Universidad de Alicante. [https://tei-c.org/Vault/P4/Lite/teiu5\\_sp.html](https://tei-c.org/Vault/P4/Lite/teiu5_sp.html)
- Iniciativa de decreto por el que se expide la Ley de Depósito Legal., 2002, H. Congreso de la Unión de México, 24 septiembre 2002, Constitución Política de los Estados Unidos Mexicanos (2002). [http://sil.gobernacion.gob.mx/Archivos/Documentos/2002/09/asun\\_61565\\_20020924\\_845445.pdf](http://sil.gobernacion.gob.mx/Archivos/Documentos/2002/09/asun_61565_20020924_845445.pdf)
- Reforma de Decreto de Depósito Legal, 2017, Cámara de Diputados, Sesión Ordinaria (2017). <http://sil.gobernacion.gob.mx/Reportes/Sesion/reporte.php?cveSesion=3631682>

- Ley General de Bibliotecas, CS-LXIV-III-1P-016, Cámara de Diputados, Ley General de Bibliotecas (2020). <http://gaceta.diputados.gob.mx/Gaceta/64/2020/dic/20201201-I.html#Minuta3>
- Cameron, L. (2018). What to Know About the Scientist Who Invented the Term "Software Engineering [First Software Engineer | IEEE Computer Society]. *Events*. <https://www.computer.org/publications/tech-news/events/what-to-know-about-the-scientist-who-invented-the-term-software-engineering/>
- Candás-Romero, J. (2006). *El papel de los metadatos en la preservación digital* (Journal Article N.º 2). El Profesional de La Información. <http://eprints.rclis.org/8359/>
- Cassini, A. (2016). Modelos científicos. En C. E. Vanney, *Diccionario Interdisciplinar Austral* (Web). Universidad Austral. [http://dia.austral.edu.ar/Modelos\\_cient%C3%ADficos](http://dia.austral.edu.ar/Modelos_cient%C3%ADficos)
- CCSDS. (2012). *Reference Model for an Open Archival Information System (OAIS)*. 135.
- CERLALC-UNESCO, C. R. para el F. del L. en A. L. y el C., UNESCO. (2019). *Recomendaciones para la actualización de las normas sobre el depósito legal en América Latina*. Centro Regional para el Fomento del Libro en América Latina y el Caribe, UNESCO. <https://cerlalc.org/publicaciones/recomendaciones-para-la-actualizacion-de-las-normas-sobre-deposito-legal-en-america-latina/>
- Ceruzzi, P. E., Institution), P. E. (Curator of A. E. and C. C., National Air & Space Museum/ Smithsonian, & Ceruzzi, C. of A. E. and C. P. E. (2003). *A History of Modern Computing*. MIT Press.
- Chartier, R. (2006). Materialidad del texto, textualidad del libro. *Orbis Tertius*, 9. [http://www.memoria.fahce.unlp.edu.ar/art\\_revistas/pr.201/pr.201.pdf](http://www.memoria.fahce.unlp.edu.ar/art_revistas/pr.201/pr.201.pdf)
- Chávez, V. M. (2018). El bibliotecario y la gestión administrativa [Billet]. *Amontonamos las palabras: Blog de la Biblioteca de El Colegio de México*. <https://bdcv.hypotheses.org/1680>
- De Decreto de Depósito Legal, 2013, Cámara de Diputados, 31 de octubre de 2013, Constitución Política de los Estados Unidos Mexicanos (2013). [http://sil.gobernacion.gob.mx/Archivos/Documentos/2013/11/asun\\_3043774\\_20131121\\_1383840627.pdf](http://sil.gobernacion.gob.mx/Archivos/Documentos/2013/11/asun_3043774_20131121_1383840627.pdf)
- Comisión de Cultura. (2018). *Dictámenes a Discusión y Votación* (Información parlamentaria LXIV/1PPO-68/87797; Senado de la República, p. 1). Senado de la República. [https://www.senado.gob.mx/64/gaceta\\_del\\_senado/documento/87797](https://www.senado.gob.mx/64/gaceta_del_senado/documento/87797)
- Congreso de la Unión México, C. del. (2021, mayo 5). *Noticias del Congreso—Diputados avalan expedir nueva Ley General de Bibliotecas*. Canal del Congreso. [https://www.canaldelcongreso.gob.mx/noticias/14221/Diputados\\_avalan\\_expedir\\_nueva\\_Ley\\_General\\_de\\_Bibliotecas](https://www.canaldelcongreso.gob.mx/noticias/14221/Diputados_avalan_expedir_nueva_Ley_General_de_Bibliotecas)
- Cruz Quintana, F. (2015). El libro transfigurado. *Tema y variaciones de literatura, Semestre 2*(45), 297-318. <http://zaloamati.azc.uam.mx/handle/11191/5105>
- Dappert, A., & Enders, M. (2010). Digital Preservation Metadata Standards. *Information Standards Quarterly*, 22(2), 11.
- DCMI. (2017). *DCMI: Dublin Core™*. <https://www.dublincore.org/specifications/dublin-core/>
- Díez-Borque, J. M. (1985). *El libro: De la tradición oral a la cultura impresa*. Editorial Montesinos.
- Dijkstra, E. (1972). The humble programmer. *ACM Turing*, 10, 859-866.
- DPC, D. P. C., OCLC, Lavoie, B., & Gartner, R. (2005). *Technology Watch Report Preservation Metadata* (p. 21). OCLC Online Computer Library Center Inc., Oxford University Library Services and Digital Preservation Coalition.
- EDItEUR. (2009). *ONIX*. <https://www.editeur.org/8/ONIX/>
- Estrada, M. (2017). Decreto. En *México | Enciclopedia Jurídica Online*. <https://mexico.leyderecho.org/decreto/>

- Fernández de Zamora, R. M. (2000). La Biblioteca Nacional de México hacia el nuevo siglo. *Métodos de Información (MEI)*, 7(40), 72-77. <http://hdl.handle.net/10760/4504>
- Finkelstein, D., McCleery, A., & Finkelstein, H. of the C. for O. L. D. (Eds.). (2006). *The Book History Reader* (segunda). Psychology Press.
- Galina Russell, I., Morfin, J., & Ramírez-Molina, A. (2019). E-legal Deposit at the Biblioteca Nacional de México (National Library of Mexico). En P. Gooding & M. Terras (Eds.), *Electronic Legal Deposit: Shaping the Library Collections of the Future* (pp. 57-56). Facet publishing. doi:10.29085/9781783303786.005
- Gartner, R., & Lavoie, B. (2013). *Preservation Metadata (2nd Edition)* (2nd ed.). Digital Preservation Coalition. <https://doi.org/10.7207/twr13-03>
- G&G. (2020). *ESTUDIO SECTOR EDITORIAL EN MÉXICO*. Gestoría en Comercio Exterior y la Oficina Comercial de Chile en Guadalajara. [https://issuu.com/prochile3.0/docs/estudio\\_editorial\\_prochile](https://issuu.com/prochile3.0/docs/estudio_editorial_prochile)
- Gil, M., & Rodríguez, J. (2011). *El paradigma digital y sostenible del libro*. Trama Editorial. <https://play.google.com/books/reader?id=M1yqjyVe9m4C&pg=GBS.PP1>
- González-Blanco, E. (2016). La edición digital de textos literarios: Planteamientos y perspectivas de futuro. *Rilce. Revista de Filología Hispánica*, 33(1), 239-258. <https://doi.org/10.15581/008.33.1.239-58>
- Guasch, P. A. (2009). *Modelado y simulación. Aplicación a procesos logísticos de fabricación y servicios* (Segunda, Vol. 118). Univ. Politèc. de Catalunya.
- Harvey, R. (2015). *The Last Decade of Digital Preservation: A Personal View from Australia*. <https://doi.org/10.1515/pdte-2015-0002>
- Henke, H. (2012). *Electronic Books and ePublishing: A Practical Guide for Authors*. Springer Science & Business Media.
- Hernández S., R. (2014). *Metodología de la Investigación* (6a ed.). McGraw Hill.
- Houston, T. (2012, mayo 24). *Dataland: The MIT's '70s media room concept that influenced the Mac*. The Verge. <https://www.theverge.com/2012/5/24/3040959/dataland-mits-70s-media-room-concept-that-influenced-the-mac>
- Howe, D. (2010). Metadata | FOLDOC. En *FOLDOC Free On-Line Dictionary of Computing*. FOLDOC. <http://foldoc.org/metadata>
- InterPARES. (2017). *Un marco de referencia para la preservación digital* (A. Barnard, A. Delgado, & J. Voutssás, Trads.; 1a ed.). Archivo General de la Nación.
- InterPARES 2. (2021). Digital object. En *Terminology Database*. International Research on Permanent Authentic Records in Electronic System. [http://www.interpares.org/ip2/ip2\\_terminology\\_db.cfm](http://www.interpares.org/ip2/ip2_terminology_db.cfm)
- InterPARES 3. (2012). Preservación digital | Multilingual Archival Terminology Database. En *Multilingual Archival Terminology Database* (TEAM México). <http://www.ciscra.org/mat/mat/term/3062>
- IPA, I. P. A. (2014). *Digital Legal Deposit and IPA special report*. International publishers Association. <https://www.internationalpublishers.org/images/news/2014/digital-legal-deposit-2014.pdf>
- Jacobson, I., Ng, P.-W., & McMahan, P. E. (2013). La Esencia de la Ingeniería de Software: El Núcleo de Semat (J. C. M. Zapata, Trad.). *Revista Latinoamericana de Ingeniería de Software*, 1(3), 71-78. <https://doi.org/10.18294/relais.2013.71-78>
- Jacobson, I., Ng, P.-W., McMahan, P., Spence, I., & Lidman, S. (2012). The Essence of Software Engineering: The SEMAT Kernel: A thinking framework in the form of an actionable kernel. *Queue*, 10(10), 40-51. <https://doi.org/10.1145/2381996.2389616>

- Jaubert, P. (2014). Depósito Legal Electrónico, El. *Revista Casa del Tiempo*, 1(Junio), 59-61.  
[http://www.uam.mx/difusion/casadel tiempo/05\\_i\\_jun\\_2014/casa\\_del\\_tiempo\\_eV\\_num\\_5\\_59\\_61.pdf](http://www.uam.mx/difusion/casadel tiempo/05_i_jun_2014/casa_del_tiempo_eV_num_5_59_61.pdf). <http://www.uam.mx/difusion/casadel tiempo>
- Khalid, A. (2014). Text Books: Ebook Vs. Print. *Journal of Education and Human Development*, 3, 16. [http://jehdnet.com/journals/jehd/Vol\\_3\\_No\\_2\\_June\\_2014/13.pdf](http://jehdnet.com/journals/jehd/Vol_3_No_2_June_2014/13.pdf)
- Kirchhoff, A., & Morrissey, S. (2014). *Preserving eBooks*. Digital Preservation Coalition.  
<https://doi.org/10.7207/twr14-01>
- Kloss Fernández del Castillo, G. (2020). La crisis del campo editorial mexicano y el imaginario de sus trabajadores. *Bibliographica*, 3(1), 13.  
<https://doi.org/10.22201/iib.2594178xe.2020.1.65>
- Laporte, C. Y., & April, A. (2018). *Software Quality Assurance*. John Wiley & Sons, Inc.  
<https://doi.org/10.1002/9781119312451>
- Larivière, J. (2000). *Legislación sobre depósito legal: Directrices*. Biblioteca de Derecho, Universidad de Ottawa, Canadá, Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. <http://www.unesco.org/new/es/communication-and-information/resources/publications-and-communication-materials/publications/full-list/guidelines-for-legal-deposit-legislation/>
- Lebert, M. (2010). *Booknología: El Libro digital (1971-2010)*. Project Gutenberg.  
<https://www.gutenberg.org/ebooks/33461>
- Letralia. (2021, julio 7). *Obligación de consignar ejemplares digitales en las bibliotecas pone en guardia a editores mexicanos*. Letralia, Tierra de Letras.  
<https://letralia.com/noticias/2021/07/07/editores-mexicanos-deposito-legal-amparo/>
- Library of Congress, E. (2016). *About—Digital Preservation*.  
<https://www.digitalpreservation.gov/about/>
- Library of Congress, E. (2019, febrero 20). *Mobipocket File Format* [Web page]. Sustainability of Digital Formats: Planning for Library of Congress Collections.  
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000472.shtml>
- LLC-EUA, T. L. L. of C. E. (2018). *Digital Legal Deposit in Selected Jurisdictions*. Library of Congress EUA, Global Legal Research Center. <https://www.loc.gov/law/help/digital-legal-deposit/>
- Luan, F., Nygård, M., & Mestl, T. (2010). A survey of digital preservation strategies. *World Digital Libraries - An International Journal*, 3(2), 133-150. <https://doi.org/10.3233/WDL-120065>
- MARC21-LC. (2008). *MARC 21 Format for Bibliographic Data: Introduction (Network Development and MARC Standards Office, Library of Congress)*.  
<https://www.loc.gov/marc/bibliographic/bdintro.html>
- METS-LC. (2016). *METS: introducción y tutorial: Sitio web oficial Metadata Encoding and Transmission Standard (METS)*.  
[https://www.loc.gov/standards/mets/METSOverview\\_spa.html](https://www.loc.gov/standards/mets/METSOverview_spa.html)
- Meyer, B. (2013, abril 4). The origin of «software engineering». *Bertrand Meyer's Technology+ Blog*. <https://bertrandmeyer.com/2013/04/04/the-origin-of-software-engineering/>
- Mintzberg, H., Brian, Q. J., & Voyer, J. (1997). *El proceso estratégico conceptos, conceptos y casos* (1a ed.). Pearson Educación.
- MODS-LC. (2020). *Metadata Object Description Schema: MODS (Library of Congress Standards)* [Webpage]. <https://www.loc.gov/standards/mods/>
- Münch, G. L. (2005). *Fundamentos de administración* (6a ed.). Trillas.
- Murray, P. R., & Squires, C. (2013). The digital publishing communications circuit. *Book 2.0*, 3(1), 3-23. [https://doi.org/10.1386/btwo.3.1.3\\_1](https://doi.org/10.1386/btwo.3.1.3_1)
- Naur, P., & Randell, B. (1969). *NATO SOFTWARE ENGINEERING CONFERENCE 1968* (p. 136) [Report on a conference sponsored]. NATO Science Committee.

- NDSA - DLF. (2019). *Levels of Digital Preservation*. National Digital Stewardship Alliance - Digital Library Federation. <http://ndsa.org//publications/levels-of-digital-preservation/>
- OCLC, & Lavoie, B. (2014). *The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition) (Second)*. Digital Preservation Coalition. <https://doi.org/10.7207/twr14-02>
- OMG, O. M. G. (2015). *Essence – Kernel and Language for Software Engineering Methods*. OMG. <http://www.omg.org/spec/Essence/1.1>
- Ordoñez, C., & Galina Russell, I. (2007). *Introducción a la edición digital (1a ed.)*. Universidad Nacional Autónoma de México (Biblioteca del Editor) / Coordinación de Difusión Cultural [UNAM] / Dirección General de Publicaciones [UNAM].
- Owens, T. (2018). *The theory and craft of digital preservation (primera)*. Johns Hopkins University Press.
- Piedras, E. (2013). *Desempeño económico de la industria editorial en México*. Nomismae. [https://www.wipo.int/copyright/es/creative\\_industries/pdf/publishing\\_mexico.pdf](https://www.wipo.int/copyright/es/creative_industries/pdf/publishing_mexico.pdf)
- Pomerantz, J. (2015). *Metadata*. The MIT Press.
- PREMIS V3-LC. (2015). *PREMIS Data Dictionary for Preservation Metadata, Version 3.0*. PREMIS Editorial Committee.
- Priani Saisó, E., & Galina Russell, I. (2019). *Una mirada al libro electrónico (1a ed.)*. Universidad Nacional Autónoma de México. <http://ru.ffyl.unam.mx//handle/10391/5405>
- Quintero, J. (2006). *La cadena de valor: Una herramienta del pensamiento estratégico*. 8(3), 14.
- RAE, & ASALE, R.-. (2020). Modelo | Diccionario de la lengua española. En «*Diccionario de la lengua española*»—*Edición del Tricentenario (Vigesimotercera)*. <https://dle.rae.es/modelo>
- Randell, B. (1996). *NATO Software Engineering Conference 1968* [Department of Computing Science University of Newcastle upon Tyne]. The NATO Software Engineering Conferences. <http://homepages.cs.ncl.ac.uk/brian.randell/NATO/NATOREports/>
- Rayl, A. J. S. (2008). NASA - NASA Engineers and Scientists-Transforming Dreams Into Reality. *50th Magazine*. [https://www.nasa.gov/50th/50th\\_magazine/scientists.html](https://www.nasa.gov/50th/50th_magazine/scientists.html)
- RDA. (2020). *About RDA* | [www.rda-rsc.org](http://www.rda-rsc.org). <http://www.rda-rsc.org/content/about-rda>
- REBIUN, R. de B. U. (2009). *Preservación digital: Guía de recursos*. Bibliotecas Universitarias Españolas.
- REDC, R. E. D. D. C. (2008). Normas / Standards. *REVISTA ESPAÑOLA DE DOCUMENTACIÓN CIENTÍFICA*, 31, 273-301.
- Reilly, E. D. (2003). *Milestones in Computer Science and Information Technology*. Greenwood Publishing Group.
- Reitz, J. M. (2014). Digital Object. En *Online Dictionary for Library and Information Science (Web)*. [https://products.abc-clio.com/ODLIS/odlis\\_about.aspx#digitalobject](https://products.abc-clio.com/ODLIS/odlis_about.aspx#digitalobject)
- Rivera D., M. A. (2009). Directrices para la creación de un programa de preservación digital. *Serie Bibliotecología y Gestión de Información*, 43, 63. <http://eprints.rclis.org>
- Romano, F. (2002). *E-Books and the Challenge of Preservation* • CLIR [Rochester Institute of Technology]. CLIR- Council on Library and Information Resources. <https://www.clir.org/pubs/reports/pub106/ebooks/>
- Rosales, N. (2012). Roger Chartier: Son las prácticas de los nativos digitales las que van a decidir la supervivencia o la muerte del libro - Universidad de Chile. *DIRCO- Universidad de Chile*. <http://www.uchile.cl/noticias/87443/roger-chartier-morira-el-libro>
- Rubio, H. J. (2014). Margaret Hamilton, la pionera de la programación que llevó el Apolo a la Luna. *El País*. [https://verne.elpais.com/verne/2014/12/11/articulo/1418314336\\_993353.html](https://verne.elpais.com/verne/2014/12/11/articulo/1418314336_993353.html)
- Rumbaugh, J., Jacobson, I., & Booch, G. (2007). *Uml: El lenguaje unificado de modelado : guía del usuario* (J. J. García Molina & J. Sáez Martínez, Trans.). Addison-Wesley.

- Rus, C. (2021, marzo 10). *E Ink desvela su nueva tecnología de tinta electrónica a color: Colores más vivos y paneles más grandes*. Xataka. <https://www.xataka.com/componentes/e-ink-desvela-su-nueva-tecnologia-tinta-electronica-a-color-colores-vivos-paneles-grandes>
- Ryan, M.-L., Emerson, L., & Robertson, B. J. (2014). *The Johns Hopkins Guide to Digital Media*. JHU Press.
- Samuelson, P., Nordhaus, W., Dieck, L., & Salazar, J. (2001). *Macroeconomía: Con aplicaciones a México* (E. Rabasco & L. Toharia, Trads.; Decimosexta, Vol. 1). McGraw Hill.
- Schnapp, J. T., & Raven, J. (2020). *The Oxford Illustrated History of the Book*. Oxford University Press.
- Ley de Depósito Legal, 2006, Cámara de Senadores, LIX/3SPo-171/8844, Ley de Depósito Legal 1 (2006). [https://www.senado.gob.mx/64/gaceta\\_del\\_senado/documento/8844](https://www.senado.gob.mx/64/gaceta_del_senado/documento/8844)
- Senso, J. A., & Rosa Piñero, A. de la. (2003). El concepto de metadato: Algo más que descripción de recursos electrónicos. *Ciência da Informação*, 32(2), 95-106. <https://doi.org/10.1590/S0100-19652003000200011>
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021a). Computer program | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021b). Computer system | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021c). Maintenance | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021d). Software | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021e). Software Engineering | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- SEVOCAB, ISO/IEC, IEEE, & PMI. (2021f). Software quality | SEVOCAB. En *Software and Systems Engineering Vocabulary*. [https://pascal.computer.org/sev\\_display/index.action](https://pascal.computer.org/sev_display/index.action)
- Sicilia, M.-A. & World Scientific (Firm). (2014). *Handbook of metadata, semantics and ontologies*. World Scientific Pub. Co. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=689758>
- Sistemas DRM. (2018, enero 22). *Uso seguro y responsable de las TIC*. <https://www3.gobiernodecanarias.org/medusa/ecoescuela/seguridad/ciudadania-y-seguridad-tic/principios-legales/derechos-de-autor/sistemas-drm/>
- Sylvestre, G. (1987). *Directrices sobre bibliotecas nacionales*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. [https://unesdoc.unesco.org/ark:/48223/pf0000076173\\_spa?posInSet=23&queryId=3108c94c-7a87-463a-b147-38020769f82b](https://unesdoc.unesco.org/ark:/48223/pf0000076173_spa?posInSet=23&queryId=3108c94c-7a87-463a-b147-38020769f82b)
- Tillett, B. (2004). Un modelo conceptual del universo bibliográfico. *The library of congress cataloging distribution service washington dc*, 8.
- Tillett, B. (2010). RDA: Antecedentes y Aspectos de su Implementación: Manual para Instructor. *Library of Congress*, 40.
- Toledo Morales, P., & Sánchez, J. M. (2002). *El libro digital: Nuevos formatos de lectura*. 126-135. <http://rabida.uhu.es/dspace/handle/10272/1049>
- Ubilla, R. A. (2016). *Libros electrónicos: Actualidad y perspectivas Guía de conceptos básicos sobres publicaciones digitales*. Naciones Unidas CEPAL. <https://docplayer.es/27453850-Division-de-publicaciones-y-servicios-web-manuales-y-guias-guia-de-libros-electronicos-de-la-cepal.html>

- UNESCO. (2002). *Memoria del Mundo: Directrices (Edición revisada 2002)* (Edición Revisada 2002; p. 71). UNESCO.  
[https://unesdoc.unesco.org/ark:/48223/pf0000105132\\_spa?posInSet=2&queryId=88f26556-e50f-4a24-8697-4d7562137b6d](https://unesdoc.unesco.org/ark:/48223/pf0000105132_spa?posInSet=2&queryId=88f26556-e50f-4a24-8697-4d7562137b6d)
- UNESCO. (2003). *Proyecto de carta para la preservación del patrimonio digital*. UNESCO.  
[http://www.r020.com.ar/enlaces/ir.php?ir\\_id=665](http://www.r020.com.ar/enlaces/ir.php?ir_id=665)
- UNESCO, & BNA, B. N. de A. (2003). *Directrices para la preservación del patrimonio digital*. UNESCO. <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>
- Viana, L. (1995). *Memoria Natural y Artificial | Cap. II Historia de la Computación (Electrónica)*. Fondo de Cultura Económica, S.A. de C.V.  
[http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/088/html/sec\\_5.html](http://bibliotecadigital.ilce.edu.mx/sites/ciencia/volumen2/ciencia3/088/html/sec_5.html)
- Voutsás, J. (2006). *Bibliotecas y publicaciones digitales* (Primera Edición). Centro Universitario de Investigaciones Bibliotecológicas.  
<https://libros.metabiblioteca.org/bitstream/001/345/4/970-32-3962-5.pdf>
- Voutsás, J. (2011). Factores culturales, económicos y sociales de la preservación documental digital. *Investigaciones Bibliotecológicas Archivonomía, Bibliotecología e Información*, 25(55), 107-150. <http://rev-ib.unam.mx/ib/index.php/ib/issue/download/2665/179>
- Voutsás, J. (2014). *Qué es inter pares* [Pdf]. IFAI - 2º Seminario Internacional sobre gestión de información y transparencia, México, D. F.  
<http://iibi.unam.mx/archivistica/que%20es%20interpares%20-%20ifai.pdf>
- Voutsás, J. (2005). El Libro Electrónico. En: «Memoria del XXII Coloquio de Investigación Bibliotecológica y de la Información». En F. Martínez Arellano, J. J. Calva González, & Universidad Nacional Autónoma de México (Eds.), *Sociedad, información y bibliotecas: Memoria del XXII Coloquio de Investigación Bibliotecológica y de la información* (pp. 62-78). Universidad Nacional Autónoma de México.
- W3C. (2021). *Publishing@W3C*. <https://www.w3.org/publishing/>
- W3C, I. G. (2017). *Web Publications for the Open Web Platform: Vision And Technical Challenges*. W3C Interest Group Note. <https://www.w3.org/TR/pwp/>
- Wang, Y. (2008). *Software Engineering Foundations: A Software Science Perspective*. CRC Press.
- Wang, Y. (2002a). On cognitive informatics. *Proceedings First IEEE International Conference on Cognitive Informatics*, 34-42. <https://doi.org/10.1109/COGINF.2002.1039280>
- Wang, Y. (2002b). On the informatics laws of software. *Proceedings First IEEE International Conference on Cognitive Informatics*, 132-141.  
<https://doi.org/10.1109/COGINF.2002.1039291>
- Waters, D. J., Kenney, A. R., Personius, L. K., Lynn, M. S., & Abell, M. D. (1994). *The Digital Preservation Consortium Mission and Goals*.  
<https://cool.culturalheritage.org/byauth/waters/dpcmiss.html>
- Wilson, T. C. (2017). Rethinking digital preservation: Definitions, models, and requirements. *Digital Library Perspectives*, 33(2), 128-136. <https://doi.org/10.1108/DLP-08-2016-0029>
- Yankelovich, N., Megrowitz, N., & Van Dam, A. (1985). Reading and writing the electronic book. *Microelectronics Journal*, 18(1), 46-46. [https://doi.org/10.1016/S0026-2692\(86\)80020-9](https://doi.org/10.1016/S0026-2692(86)80020-9)
- Zarndt, F., Carner, D., McCain, E., Clausen, T., & Wyber, S. (2018). *Results of the 2017 Survey of Electronic Legal Deposit Policies and Practices at National Libraries*. IFLA, IIPC, the Donald W. Reynolds Journalism Institute, and the University of Missouri Libraries.  
<https://www.ifla.org/node/61887>

## ANEXOS.

### I. ÍNDICE DE IMÁGENES E ILUSTRACIONES

Figura 1. Método Cualitativo.....	9
Figura 2. Diseño sistemático para el análisis.....	11
Figura 3. Visión del mundo en el Modelo IME. ....	35
Figura 4. Componentes de objeto digital, libro electrónico y software.....	37
Figura 5. Modelo OAIS.....	60
Figura 6. Jerarquía de necesidades de preservación digital. ....	61
Figura 7. Modelo de Información OAIS. ....	69
Figura 8. Estructura completa del PDI en OAIS. ....	72
Figura 9. Modelo PREMIS. ....	73
Figura 10. Vista conceptual entre las subcategorías que componen al objeto. ....	74
Figura 11. Entidades grupo1. Requerimientos Funcionales para Registros Bibliográficos .....	79
Figura 12. Entidades grupo2. Requerimientos Funcionales para Registros Bibliográficos.....	80
Figura 13. Niveles de organización BIBFRAME 2.0.....	81
Figura 14. Circuito de comunicación del libro.....	85
Figura 15. Digital Publishing Communications Circuit. ....	86
Figura 16. Digitalización Estandarizada.....	90
Figura 17. Elementos del proceso de software.....	94
Figura 18. “The Kernel Alphas” o alphas del núcleo. ....	95
Figura 19. “The Kernel Activity Spaces” o espacios de actividad del núcleo. ....	97
Figura 20. “The Kernel Competencies” o competencias del núcleo. ....	98
Figura 21. Digitalización futura. ....	100
Figura 22. Workflow de ficheros. ....	101
Figura 23. Actividades en común del proceso editorial de CEPAL. ....	103
Figura 24. Unidad gráfica editorial de CEPAL.....	103
Figura 25. Secuencia de elaboración de documentos del proceso editorial de CEPAL. ....	104
Figura 26. Diagrama de actividades en UML del proceso editorial resultado del análisis.....	118
Figura 27. Control y seguimiento, unidad de trabajo del proceso editorial propuesto.....	119
Figura 28. Cambios en unidad de trabajo control editorial del proceso propuesto .....	124
Figura 29. Edición, unidad de trabajo del proceso editorial propuesto.....	125
Figura 30. Cambios en la unidad de trabajo Producción del proceso editorial propuesto.....	128
Figura 31. Programación. Unidad de trabajo del proceso editorial propuesto. ....	129
Figura 32. Cambios en la unidad de trabajo producción del proceso editorial propuesto.....	133
Figura 33. Cambios en la unidad de trabajo control de calidad del software en la edición. ....	134
Figura 34. Distribución. Unidad de trabajo del proceso editorial propuesto. ....	135
Figura 35. Cambios en unidad de trabajo distribución del proceso editorial propuesto. ....	138
Figura 36. Proceso editorial con adecuaciones derivadas de aplicar IS.....	140
Figura 37. Modelo, nivel 1, diagrama de componentes para especificar a la publicación. ....	147
Figura 38. Modelo, nivel 2, diagrama de clases para representar involucrados en la producción de la publicación. ....	149

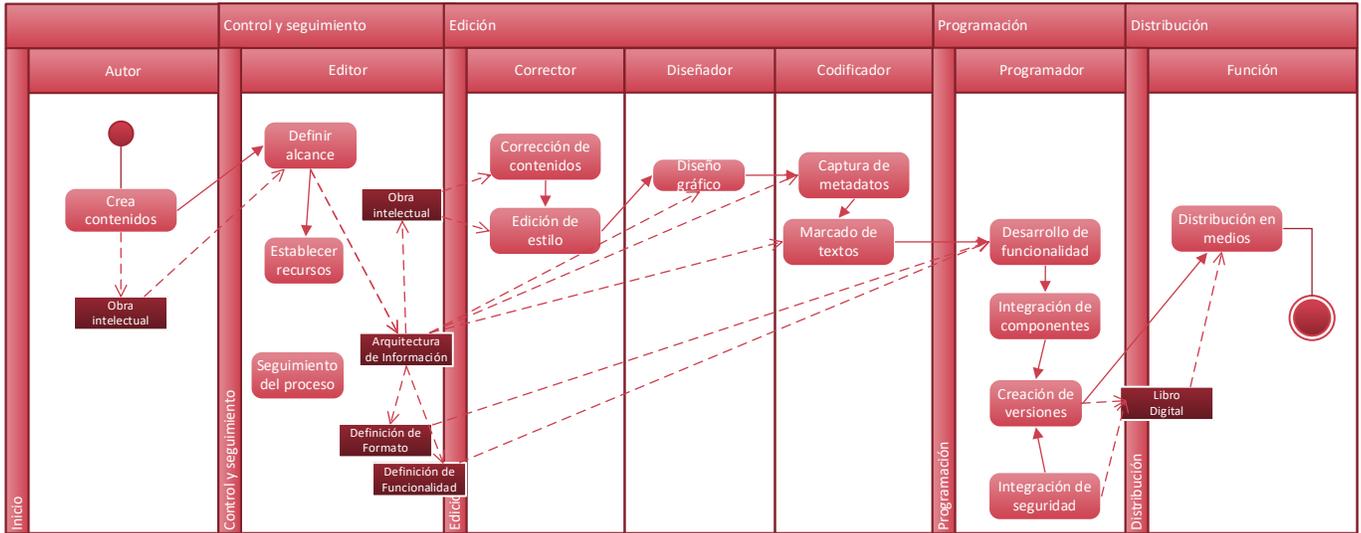
Figura 39. Modelo, nivel 3, diagrama de clases para representar involucrados en el alcance de la publicación. ....	151
Figura 40. Ambiente OAIS. ....	153
Figura 41. Modelo, nivel 4, paquetes de clases para la identificación de Ente productor y Ente regulador. ....	154
Figura 42. (AIP) Paquete de Información de Archivo OAIS. ....	156
Figura 43. Modelo, nivel 5, componentes empaquetados del Contenido Informativo (CI) de la publicación. ....	157
Figura 44. Modelo, nivel 6, diagrama de paquetes que muestra los atributos de un objeto digital. ....	158
Figura 45. Modelo, nivel 7, diagrama de casos de uso para representar el proceso editorial y sus actores. ....	160
Figura 46. Modelo, nivel 8, diagrama de paquetes para representar los factores que intervienen en el proceso editorial. ....	161

## II. ÍNDICE DE TABLAS

Tabla 1. Organizaciones interesadas en la preservación digital a nivel mundial, identificadas durante el desarrollo de esta investigación. ....	55
---	----

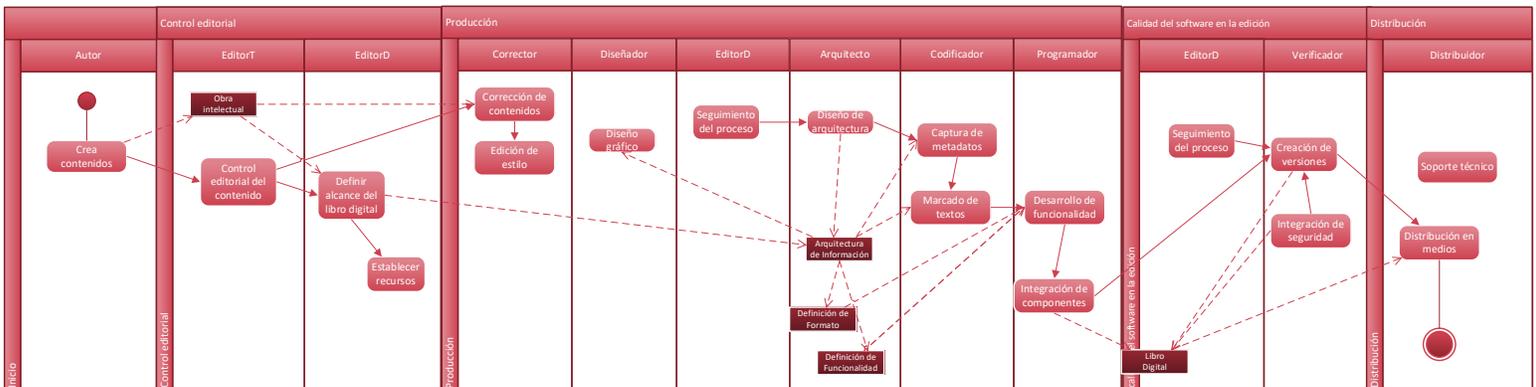
III. DIAGRAMAS DEL PROCESO EDITORIAL ELABORADO.

Diagrama de actividades en UML del proceso editorial resultado del análisis.



Fuente: La autora diciembre 2019.

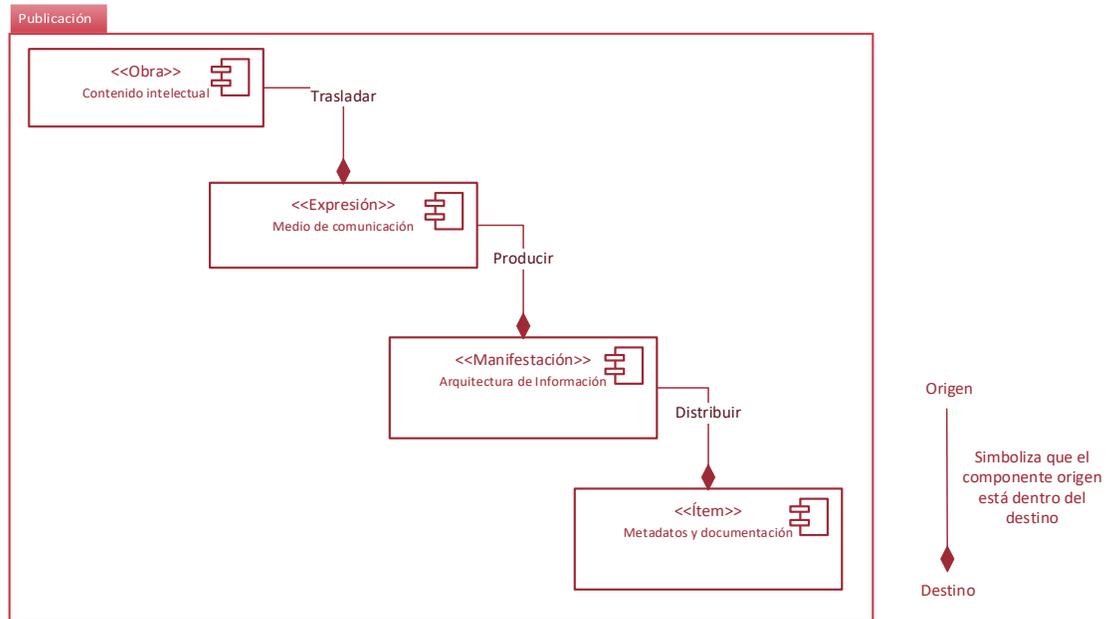
Proceso editorial con adecuaciones derivadas de aplicar IS.



Fuente: La autora diciembre 2019.

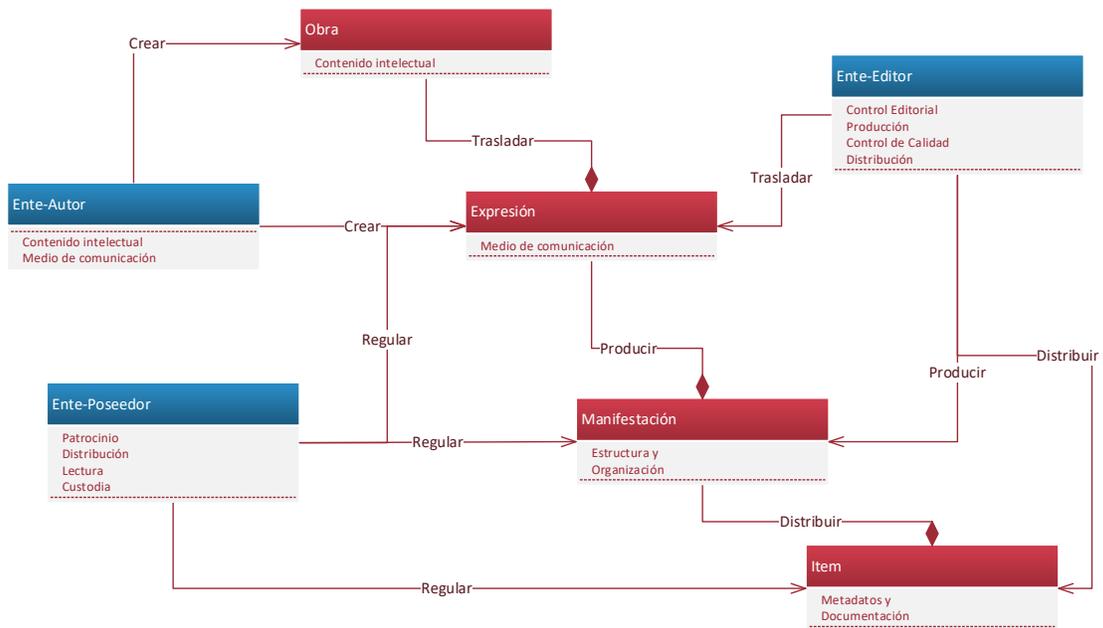
## IV. EL MODELO

Modelo, nivel 1, diagrama de componentes para especificar a la publicación.



Fuente: La autora diciembre 2019.

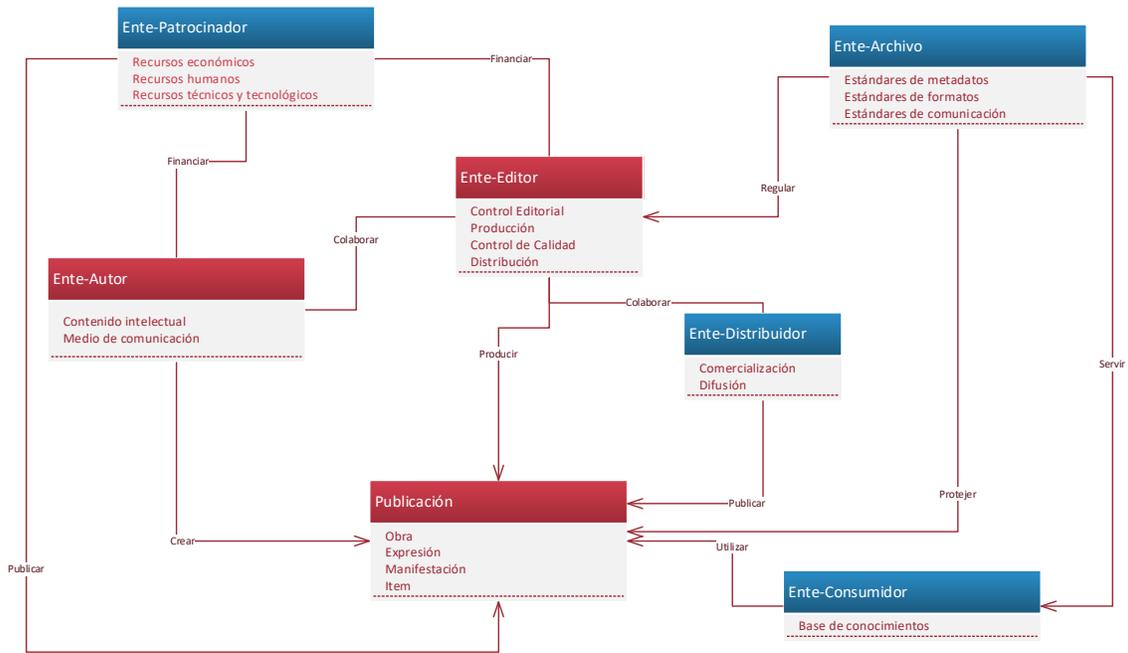
Modelo, nivel 2, diagrama de clases para representar involucrados en la producción de la publicación.



Fuente: La autora diciembre 2019.

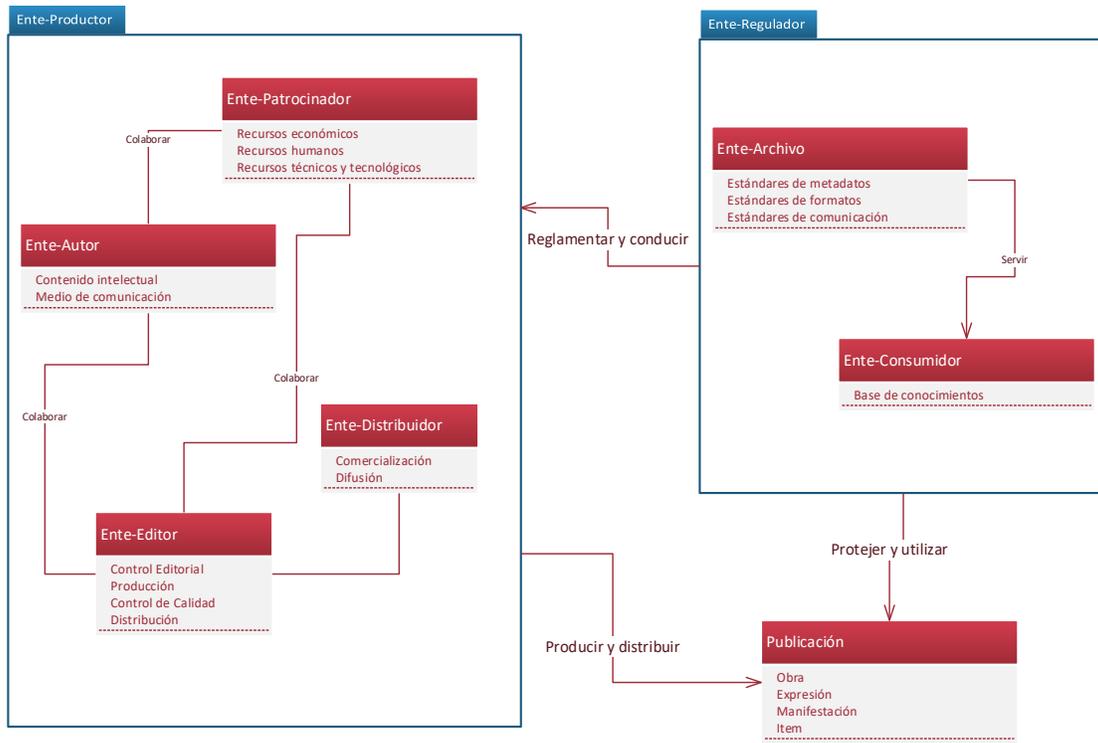
# MODELO DE INTEGRACIÓN DE METADATOS PARA PRESERVACIÓN EN LA EDICIÓN DEL LIBRO DIGITAL

Modelo, nivel 3, diagrama de clases para representar involucrados en el alcance de la publicación



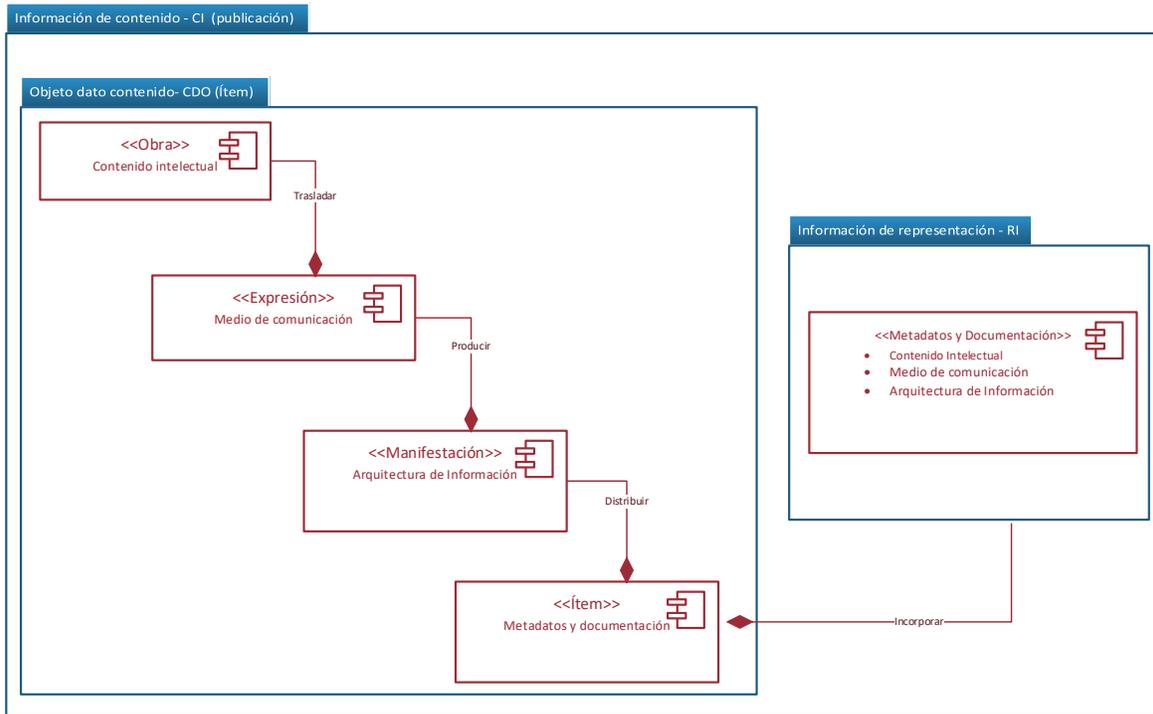
Fuente: La autora diciembre 2019.

Modelo, nivel 4, paquetes de clases para la identificación de Ente productor y Ente regulador



Fuente: La autora diciembre 2019.

Modelo, nivel 5, componentes empaquetados del Contenido Informativo (CI) de la publicación.



Fuente: La autora diciembre 2019.

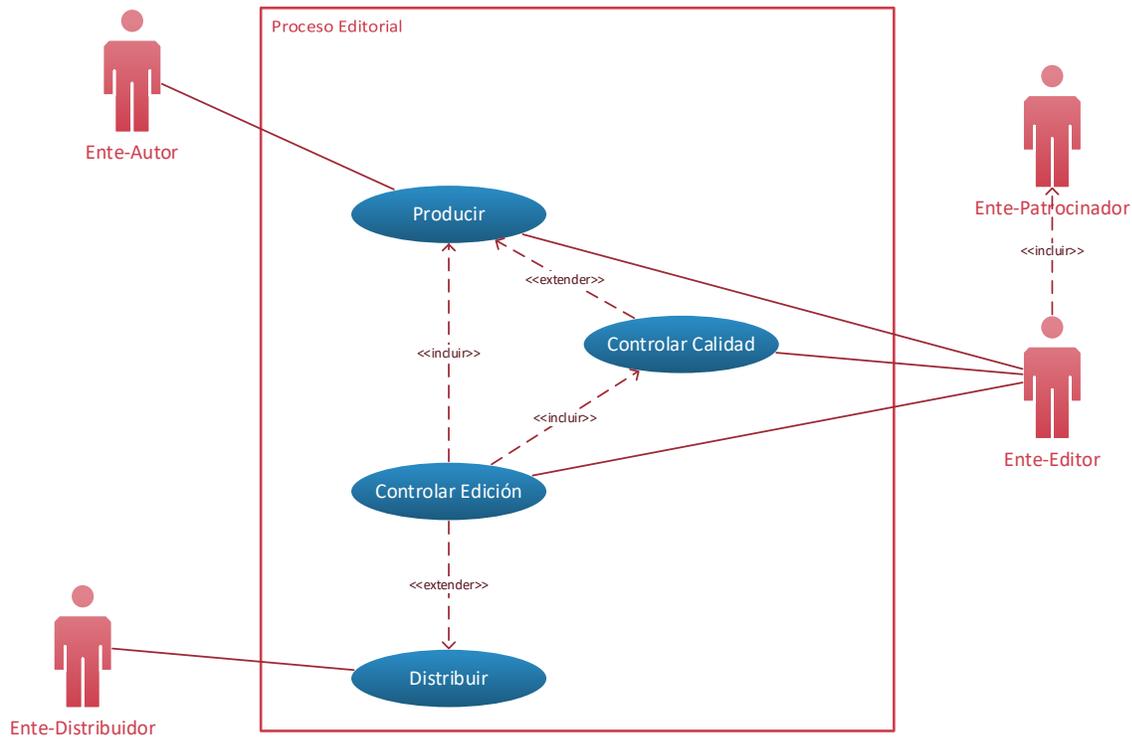
Modelo, nivel 6, diagrama de paquetes que muestra los atributos de un objeto digital



Fuente: La autora diciembre 2019.

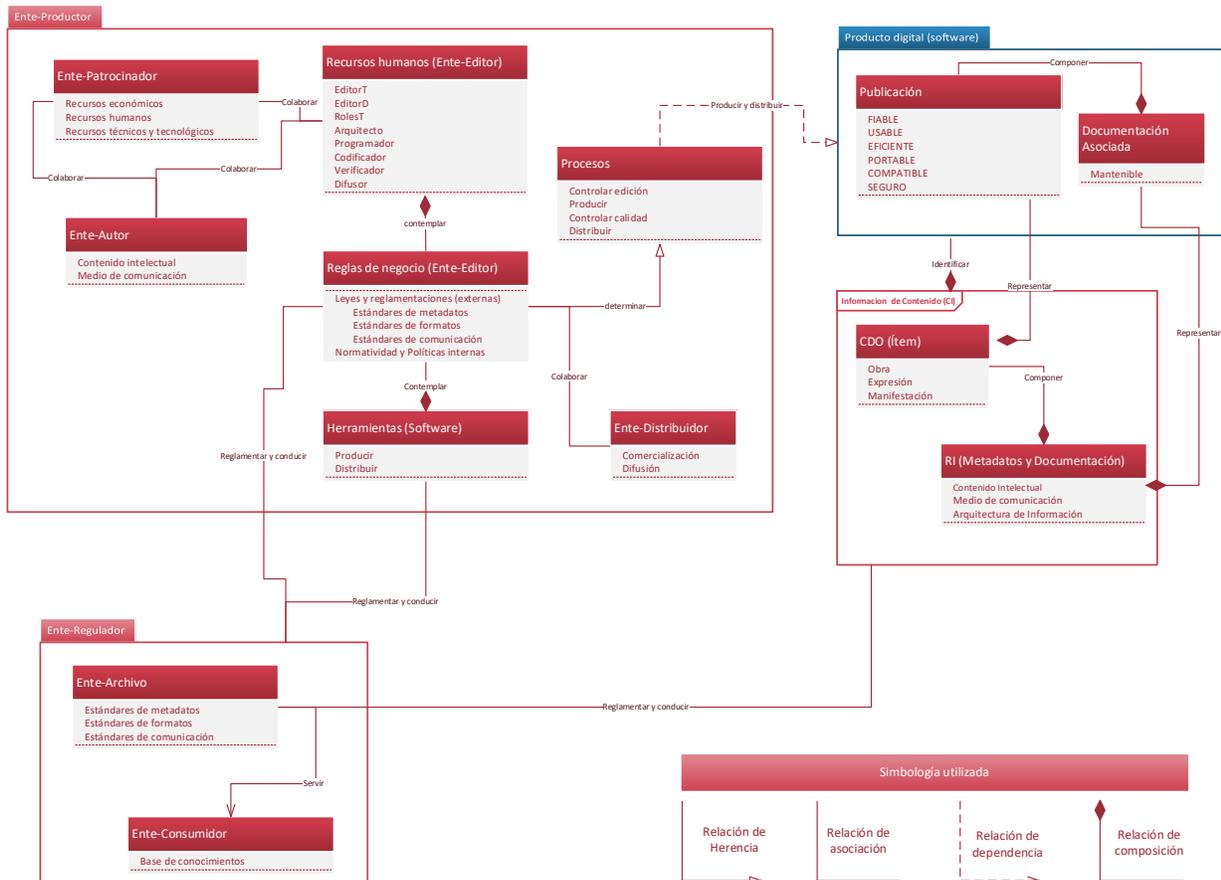
# MODELO DE INTEGRACIÓN DE METADATOS PARA PRESERVACIÓN EN LA EDICIÓN DEL LIBRO DIGITAL

Modelo, nivel 7, diagrama de casos de uso para representar el proceso editorial y sus actores



Fuente: La autora diciembre 2019.

Modelo, nivel 8, diagrama de paquetes para representar los factores que intervienen en el proceso editorial



Fuente: La autora diciembre 2019.