

Directrices
UNESCO/PERSIST sobre
selección del patrimonio
digital para su conservación a
largo plazo

Segunda edición

Mayo de 2021

Índice

| | |
|--|----|
| Prefacio | 3 |
| Introducción | 5 |
| Destinatarios..... | 5 |
| Principales retos | 5 |
| Diversidad e inclusión | 6 |
| Estructura de las Directrices | 6 |
| Parte 1: El impacto de la preservación digital en las decisiones relativas a la selección | 7 |
| Definición de preservación digital..... | 7 |
| ¿En qué se diferencia la recopilación de material digital de la recopilación de material tradicional?..... | 8 |
| Tipos de material digital..... | 8 |
| Elaboración de los criterios de selección..... | 9 |
| Criterios de selección | 9 |
| Estrategias de recopilación del patrimonio digital | 10 |
| Selección | 10 |
| Recopilación exhaustiva | 11 |
| Muestreo representativo | 11 |
| Considerar el contexto | 11 |
| Árbol de decisión para la selección en una institución concreta..... | 12 |
| Paso 1: Identificación | 12 |
| Paso 2: Marco jurídico | 13 |
| Paso 3: Aplicación de los criterios de selección | 13 |
| Parte 2: Análisis detallado de la recopilación en el entorno digital | 17 |
| Selección digital en bibliotecas, museos y archivos | 17 |
| Bibliotecas..... | 18 |
| Museos..... | 18 |
| Archivos | 19 |
| Alianzas para la selección y preservación digital..... | 20 |
| Cooperación internacional..... | 20 |
| El papel de las instituciones nacionales..... | 20 |
| La importancia de la promoción..... | 21 |
| La influencia del entorno legal..... | 22 |
| Otras consideraciones legales: derechos de autor y privacidad | 23 |
| Conclusión | 25 |

| | |
|--|----|
| Apéndice 1: Coleccionismo digital frente a coleccionismo tradicional..... | 26 |
| Similitudes | 26 |
| Diferencias..... | 26 |
| Apéndice 2: Análisis detallado del código fuente del <i>software</i> | 28 |
| Apéndice 3: Análisis detallado de los datos de investigación | 29 |
| Preguntas esenciales para la selección | 29 |
| Apéndice 4: Análisis detallado de la evaluación y selección de los medios sociales | 31 |
| Apéndice 5: Análisis detallado de la inteligencia artificial | 33 |
| Introducción..... | 33 |
| Definición | 33 |
| ¿Cómo se evalúa y selecciona la IA? | 34 |
| Otras consideraciones | 37 |
| Apéndice 6: Gestión de los metadatos..... | 40 |
| Metadatos | 40 |
| Almacenamiento de los metadatos | 41 |
| Meta metadatos | 41 |
| Apéndice 7: Referencias | 42 |

Prefacio

El mundo digital se mueve a la velocidad del rayo. Tanto los contenidos como las tecnologías digitales experimentan importantes avances a diario, lo que supone un gran desafío para las instituciones patrimoniales y otras organizaciones que se ocupan de la información a la hora de seguir seleccionando, preservando y accediendo a la producción del patrimonio documental del mundo. Para garantizar la accesibilidad a largo plazo al patrimonio digital más importante, son fundamental las intervenciones tempranas de identificación y preservación.

Han pasado cinco años desde que se elaboró la primera edición de las "Directrices UNESCO/PERSIST sobre selección del patrimonio digital para su conservación a largo plazo" con el fin de apoyar a las instituciones patrimoniales y de investigación en el proceso de selección del patrimonio digital para su conservación sostenible a largo plazo. En un mundo digital, cinco años son una eternidad, por lo que se hizo necesaria una segunda edición de las Directrices.

La primera edición de las Directrices se basó en la labor desarrollada en el marco del Proyecto UNESCO/PERSIST, que surgió de la Conferencia de la Memoria del Mundo celebrada en Vancouver (Canadá) en 2012. El Proyecto PERSIST es una iniciativa de colaboración entre la UNESCO, la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA), el Consejo Internacional de Archivos (ICA) y otros asociados. En 2020, el proyecto PERSIST se integró en los trabajos del Subcomité de Preservación del Programa Memoria del Mundo de la UNESCO.

Repercusión

La primera edición de las Directrices se publicó en nueve idiomas en formato impreso y electrónico, y se descargó cientos de veces de los sitios web de la UNESCO y la IFLA. El texto fue utilizado principalmente por los profesionales de la información para adoptar decisiones sobre la selección con fines de preservación. Además, en algunos casos, las Directrices se utilizaron en talleres de formación para la gestión de la información digital. Las Directrices también tenían por objetivo concienciar a los gobiernos y a la sociedad civil sobre la naturaleza fugaz de la información digital. Se espera que, a medida que se vayan afianzando las actividades de promoción de las organizaciones que se ocupan del patrimonio documental, aumente la difusión de esta segunda edición de las Directrices.

Las tres asociaciones internacionales principales relacionadas con el patrimonio documental —la IFLA, el ICA y el Consejo Internacional de Museos (ICOM)— han hecho suyas estas Directrices y las promueven en el marco de sus circunscripciones y redes.

Relación con la UNESCO

La UNESCO ha apoyado firmemente las ediciones anterior y actual de las Directrices, utilizándolas como instrumento de formación para sensibilizar acerca de la preservación y el acceso a la información digital en muchas partes del mundo desarrollado y menos desarrollado. En 2015, la Conferencia General de la UNESCO hizo suya la Recomendación relativa a la Preservación del Patrimonio Documental, comprendido el Patrimonio Digital, y el Acceso al mismo. Por consiguiente, cada cuatro años, los Estados Miembros de la UNESCO deben informar sobre su aplicación de la Recomendación de 2015. Estas Directrices son un

recurso útil que puede servir de ayuda a los Estados Miembros de la UNESCO para aplicar la Recomendación de 2015.

Grupo de Revisión de la Redacción

Muchos profesionales de la información han participado en la elaboración de esta segunda edición de las Directrices, y lo han hecho en un momento de grandes restricciones debido a la pandemia de COVID-19. Han sido muy rigurosos en el desempeño de esta tarea y merecen toda nuestra gratitud por haber compartido su experiencia con tanto entusiasmo.

Ingrid Parent, Presidenta (IFLA, bibliotecaria universitaria emérita, Universidad de Columbia Británica - UBC)

Claire McGuire, Secretaria (Responsable de Política e Investigación, IFLA)

Anthea Seles (Secretaria General del Consejo Internacional de Archivos)

Davide Storti (Especialista del Programa, Sector de Comunicación e Información, UNESCO)

Fackson Banda (Especialista del Programa, Programa Memoria del Mundo de la UNESCO) (*ex officio*)

Frédéric Blin (Jefe de Servicios y Colecciones, Biblioteca Nacional y Universitaria de Estrasburgo)

Gordon McKenna (Director de Normas, Collections Trust) (desde 2021)

Ivy Lee (Directora de Funciones Estatutarias e Investigación, Biblioteca Nacional de Singapur) (desde 2020)

Jenna Murdock Smith (Archivista Principal, Rama de Archivos, Biblioteca y Archivos del Canadá)

Julia Chee (Subdirectora de Colecciones, Biblioteca Nacional de Singapur) (hasta 2020)

Monika Hagedorn-Saupe (Presidenta del ICOM CIDOC, Fundación del Patrimonio Cultural Prusiano) (hasta 2020)

Steve Knight (Director del Programa de Investigación de la Preservación, Biblioteca Nacional de Nueva Zelanda)

Winston Roberts (asesor comercial principal, Oficina del Bibliotecario Nacional, Biblioteca Nacional de Nueva Zelanda)

Esperamos que estas Directrices le resulten útiles para dar respuesta a sus necesidades concretas en materia de preservación digital.

Sírvase enviar sus comentarios o preguntas a la IFLA: ifla@ifla.org

Introducción

Tradicionalmente, las instituciones que se ocupan del patrimonio –bibliotecas, archivos y museos– han asumido la responsabilidad de preservar los recursos intelectuales y culturales producidos por el conjunto de la sociedad. Esta importante misión está actualmente en riesgo en todo el mundo debido al ingente volumen de información producida y compartida a diario en formato digital. En 2020, el Global DataSphere de International Data Corporation (IDC), que mide la cantidad de datos creados y consumidos en el mundo cada año, predijo que la cantidad de datos creados entre 2020 y 2023 superará la cantidad de datos generados en los 30 años anteriores (IDC, 2020). Este crecimiento exponencial supone un reto cada vez mayor para quienes se ocupan de la conservación del patrimonio digitalizado y de origen digital para su accesibilidad a largo plazo.

Destinatarios

Estas Directrices fueron concebidas para su aplicación en todo el mundo y en todos los contextos. En las Directrices se reconoce que no se puede ser demasiado específico en su aplicación porque las políticas en materia de patrimonio cultural varían dependiendo de los países, regiones e instituciones. Constituyen un punto de partida para las bibliotecas, los archivos, los museos y otras instituciones que se ocupan del patrimonio a la hora de elaborar sus propias políticas de selección para la conservación.

Estas Directrices se dirigen principalmente a los profesionales de las instituciones culturales encargados de adoptar las decisiones cotidianas sobre qué materiales digitales son candidatos para su accesibilidad a largo plazo. Sirven de apoyo a los profesionales para revisar las políticas existentes, puesto que resaltan las cuestiones importantes que se deben tener en cuenta y proporcionan orientación para redactar las políticas institucionales. Los administradores de alto nivel de las instituciones patrimoniales también encontrarán útiles estas Directrices a la hora de planificar y asignar prioridades a las estrategias y los presupuestos para la gestión de sus recursos.

Principales retos

Preservar este enorme producto es difícil, no únicamente por su tamaño, sino porque buena parte del mismo es de naturaleza efímera. La información digital no tiene la misma longevidad que los objetos físicos, los documentos y los libros, que a menudo sobreviven durante siglos. Los formatos de archivo digital, los soportes de almacenamiento y los sistemas están en constante transformación y ponen en riesgo la legibilidad e integridad futuras del patrimonio digital en plazos mucho más cortos que los que afectan al deterioro del papel y los objetos físicos. Además, los sistemas digitales que son a la vez instrumentos para la creación de contenidos y creaciones en sí mismos están empezando a ser reconocidos como contenidos digitales que hay que conservar.

La supervivencia del patrimonio digital está mucho menos asegurada que la de sus equivalentes tradicionales en nuestras colecciones. La identificación del patrimonio digital importante y la intervención temprana son esenciales para asegurar su conservación a largo plazo. También lo es la gestión de la carga económica de la financiación de la preservación,

que supone un reto para los presupuestos nacionales e institucionales, además de ser fuente de desigualdad entre países y comunidades.

Está claro que la responsabilidad de la conservación del patrimonio digital va más allá de las instituciones culturales y requerirá el compromiso y la cooperación del sector público, el sector privado y los creadores de contenidos. El sector privado también se enfrentará al reto de preservar y garantizar el acceso a su propia información digital. Esto podría considerarse como una oportunidad de cooperación entre los sectores privado y público, y para llevar a cabo proyectos conjuntos.

Diversidad e inclusión

El material objeto de recopilación al que se refieren estas Directrices pretende incluir contenidos digitales creados por todos los grupos étnicos, religiosos, de género, sociales y políticos que se encuentran en todas las regiones del mundo, o acerca de ellos. El Grupo de Trabajo sobre el Contenido recomienda que los archivos, las bibliotecas y los museos consulten con las comunidades infrarrepresentadas y colaboren con ellas a la hora de tomar decisiones de selección para garantizar que el patrimonio documental creado por estas comunidades, y acerca de ellas, sea identificado para su conservación digital a largo plazo.

El Grupo de Revisión reconoce que no está en las mejores condiciones para responder a las necesidades y perspectivas de todos los grupos marginados en estas Directrices. Por lo tanto, se acogen con satisfacción y se alientan las iniciativas futuras dirigidas por esas comunidades, como las Directrices para la selección del patrimonio digital indígena para su conservación a largo plazo.¹

Estructura de las Directrices

Esta segunda edición de las Directrices está organizada en dos partes. En la primera parte, sobre el impacto de la preservación digital en las decisiones relativas a la selección, se ofrece una visión concisa de las decisiones relativas a la selección que deben tomar los profesionales y otros agentes para preservar esta nueva e importante parte del patrimonio documental mundial. En la segunda parte se ofrece un análisis más detallado de la recopilación en el entorno digital.

Estas dos partes van seguidas de varios apéndices, en los que se proporciona información más detallada en relación con varias cuestiones digitales y su contexto.

¹ Para consultar información introductoria sobre la gobernanza de los datos indígenas, recomendamos consultar los Principios CREA para la Gobernanza de Datos Indígenas: <https://static1.squarespace.com/static/5d3799de845604000199cd24/t/5f0621e031811d08f9c4cd27/1594237427481/Principios+CREA.pdf>

Parte 1: El impacto de la preservación digital en las decisiones relativas a la selección

Definición de preservación digital

Para definir la preservación digital, nos remitimos a la definición que proporciona la Coalición de la Preservación Digital:

La preservación digital se refiere a la serie de actividades gestionadas necesarias para garantizar el acceso continuo a los materiales digitales durante el tiempo que sea necesario. La preservación digital... se refiere a todas las acciones necesarias para mantener el acceso a los materiales digitales más allá de los límites de los fallos o de los cambios tecnológicos y organizativos de los medios de difusión. Estos materiales pueden ser registros creados durante el día a día de una organización; materiales concebidos como digitales con una finalidad específica (por ejemplo, recursos didácticos); o productos de los proyectos de digitalización (Manual de Preservación Digital, 2015).

La preservación digital garantiza el acceso a largo plazo a la información digital a lo largo del tiempo, las tecnologías y los cambios semánticos, y tiene beneficios a largo plazo, que puede ser sociales (prueba de gobierno), culturales (identidad nacional) y económicos (uso y reutilización, innovación).

El cambio digital ha modificado radicalmente la forma de producir, almacenar, conservar, difundir, consumir y rentabilizar el patrimonio cultural. También ha cambiado drásticamente el papel de los ciudadanos, que han pasado de ser observadores pasivos a participantes activos y productores de contenidos con numerosas, nuevas y emocionantes oportunidades de participación, uso creativo y acceso (Tammaro, 2016).

El entorno digital ha creado nuevas formas de expresión, que abarcan desde sitios web y redes sociales interactivas hasta bases de datos de investigación privadas, obras de arte digitales y plataformas de juegos en línea. Estos materiales desdibujan los límites y las estructuras de responsabilidad, lo que supone un reto frente a las antiguas formas de recopilación.

A menudo, los mandatos y las políticas de recopilación no cubren esas nuevas formas de patrimonio digital. Nuestro descuido con esas nuevas formas aumenta el riesgo de crear grandes vacíos en nuestro patrimonio cultural para las generaciones venideras.

Por ejemplo, puede ser que el valor de las entradas individuales en blogs o en redes sociales sea inapreciable, pero de manera colectiva constituyen un registro único de la sociedad contemporánea; las conversaciones, pensamientos y logros de miles de millones de personas. Si se preservaran, serían una fuente incomparable de conocimientos para las generaciones futuras. Centrarse únicamente en la "mejor" parte de este producto, dependiendo del mandato de las instituciones, introduciría sesgos e impediría cualquier análisis de la creación contemporánea digital como parte de un conjunto.

Pocas o ninguna de las instituciones culturales tienen los recursos, y, en algunos casos, los derechos, para recopilar y preservar en conjunto el producto social en formato digital. *Esta es la paradoja de la selección en la era digital*. La selección es tan esencial como económica y técnicamente imposible y, a menudo, está legalmente prohibido recopilar todo el patrimonio digital actual.

Seleccionar en aras de la conservación a largo plazo será, pues, una función fundamental de las instituciones que se ocupan del patrimonio en la era digital y tendrá que abordarse con planteamientos prácticos, pragmáticos y colaborativos para crear un registro histórico representativo.

¿En qué se diferencia la recopilación de material digital de la recopilación de material tradicional?

Los materiales digitales plantean retos únicos, ya que a menudo solo pueden archivarse adecuadamente en formato digital. No existe un equivalente no digital que aglutine toda la información esencial y la funcionalidad que proporciona, y la preservación digital solo es útil si se puede acceder al material en una fecha posterior (Coalición de la Preservación Digital, 2015).

Las principales diferencias entre la recopilación de materiales digitales y tradicionales son:

- La velocidad y la complejidad de los cambios en el ámbito digital
- La necesidad de instrumentos y servicios intermediarios para entender y manipular el objeto digital
- El volumen de información que se debe revisar para la selección
- Obsolescencia constante del *hardware* y el *software* con el paso del tiempo
- La velocidad y la casi invisibilidad del deterioro de los contenidos digitales
- La escala a la que habrá que llevar a cabo la recopilación, el almacenamiento y la preservación digitales será cada vez mayor, ya que con el tiempo se pasará de los petabytes a los exabytes.

Véase el Apéndice 1: Coleccionismo digital frente a coleccionismo tradicional para conocer con más detalle las principales similitudes y diferencias entre la recopilación tradicional y la digital.

Los profesionales también han de ser conscientes de que deben seleccionar y valorar materiales en distintos formatos, ya que los formatos impresos y digitales suelen coexistir en lo que puede denominarse una larga *cola híbrida*. Por lo tanto, el planteamiento de las prácticas de selección debe ser holístico, en lugar de ser exclusivamente digital o impreso/analógico. Se plantean otros retos relacionados con la interconexión entre los formatos impreso y digital para permitir a los investigadores una visión de conjunto, pero eso no entra en el ámbito de competencia de las presentes Directrices.

Tipos de material digital

Hoy por hoy, muchas formas tradicionales de patrimonio cultural tienen sus equivalentes digitales, que a menudo encajan bien con nuestras prácticas y mandatos actuales. Se trata de libros, publicaciones periódicas, documentos gubernamentales, correspondencia privada,

diarios personales, mapas, fotografías, grabaciones cinematográficas y sonoras, artefactos y obras de arte.

Además de los equivalentes digitales de materiales analógicos, el patrimonio digital también adopta la forma de medios sociales, material de realidad virtual (RV) y de realidad aumentada (RA), arte digital o archivos web, entre otras.

Si bien no es posible incluir una lista pormenorizada de todos los tipos de materiales digitales que se deben analizar, el Grupo de Revisión consideró importante examinar varios tipos de materiales que podrían no estar incluidos en las estrategias tradicionales de selección de contenidos.

Sírvase consultar los Apéndices para obtener mayor información sobre la selección del código fuente del *software*, los datos de investigación, los medios sociales y la inteligencia artificial para su conservación a largo plazo.

Elaboración de los criterios de selección

En el marco institucional actual, las misiones, mandatos y políticas de desarrollo de colecciones para las colecciones físicas aportarán, en la mayoría de los casos, el punto de partida y las orientaciones básicas para valorar y seleccionar el patrimonio digital. Estos elementos deben adaptarse para abarcar las nuevas formas de expresión digital.

La evaluación y valoración del patrimonio digital debe basarse en muchos de los mismos principios en los que se basa la selección tradicional, como el contexto y la procedencia, reconociendo, por otra parte, que algunos aspectos de la colección tradicional no pueden trasladarse al entorno digital. A la hora de tomar decisiones sobre la selección del patrimonio digital, también es necesario tener en cuenta las nuevas cuestiones de accesibilidad, uso y conservación a largo plazo.

Criterios de selección

Una institución debería poder contestar a estas preguntas tras valorar la **importancia** relativa del patrimonio digital en relación con su mandato y con el público; evaluar su **sostenibilidad**, es decir, la capacidad de la institución para conservar su acceso y uso a largo plazo; y tomar en consideración su **disponibilidad** en otras instituciones que se ocupan del patrimonio, es decir, la posibilidad de su conservación en otros lugares, y cuál es la institución o grupo comunitario más apropiado para su conservación.²

Los conceptos de importancia y sostenibilidad en este entorno deben evaluarse a la luz de los mandatos y recursos institucionales. Para determinar la disponibilidad hay que fijarse en otras instituciones de la comunidad patrimonial a fin de evaluar el nivel de riesgo para la supervivencia del patrimonio digital a largo plazo. Debe prestarse especial atención al patrimonio que corre el riesgo de perderse a corto y medio plazo, así como al patrimonio

² Cabe señalar que las instituciones pueden optar por seleccionar documentos de archivo importantes que pueden no ser sostenibles, con vistas a que se desarrollen infraestructuras o tecnologías que permitan su conservación o legibilidad en un momento posterior.

considerado de gran importancia para la humanidad, como las colecciones incluidas en el Registro de la Memoria del Mundo de la UNESCO.

Habida cuenta de que cada institución de la memoria es única en cuanto a su mandato, política de colecciones y recursos, a continuación, ofrecemos una serie de medidas y preguntas que pueden servir para adoptar decisiones relativas a la selección digital. Este planteamiento puede adaptarse a las necesidades individuales de las instituciones, que tienen objetivos y tamaños diversos. Incluso si no se adoptan, estas medidas pueden constituir el punto de partida de un debate institucional sobre la selección del patrimonio digital para su conservación a largo plazo.

Estrategias de recopilación del patrimonio digital

Los planteamientos actuales de las instituciones del patrimonio deben evolucionar para adaptarse al entorno digital. Es probable que esto incluya la adaptación de una o más de las siguientes estrategias o enfoques.

Selección

La selección se produce cuando los profesionales del patrimonio — archiveros, bibliotecarios y conservadores— identifican los materiales que deben añadirse a sus colecciones basándose en criterios específicos. Estos criterios pueden variar ampliamente en función del tipo de institución, su mandato de recopilación, sus recursos y el tipo y la cuantía del material disponible para su adquisición. Los criterios de selección suelen estar expresados o definidos en una política de recopilación o adquisición, y pueden basarse en los criterios siguientes (que también pueden combinarse):

- *Función.* Las instituciones que se ocupan del patrimonio documental, como los archivos, seleccionarán los documentos con valor a largo plazo sobre la base de la labor realizada por una institución o gobierno. Por ejemplo, los telegramas diplomáticos desempeñan una importante función de información en los Ministerios de Relaciones Exteriores.
- *Asunto/tema.* Una institución se centrará en una o varias áreas temáticas, que tratará de documentar. Por ejemplo, todos los sitios web dedicados a un pintor o localidad concretos, o un rastreo de la web para documentar un acontecimiento específico, como unas elecciones políticas o un festival artístico.
- *Creador/Procedencia.* Una institución se centrará en determinados creadores de patrimonio o en su procedencia. Por ejemplo, un archivo puede adquirir los registros digitales de los autores de una determinada región; o un museo puede coleccionar obras de los artistas de un determinado movimiento.
- *Tipo/Formato.* Una institución puede realizar la recopilación en función del tipo o formato del contenido (por ejemplo, fotografía digital, grabaciones musicales, películas, videojuegos).

En algunos casos, las instituciones pueden optar por capturar todo el material del patrimonio digital ahora y aplicar los criterios de selección en una fecha posterior, como forma de

selección diferida. A continuación, se describen dos prácticas específicas de recopilación: la recopilación exhaustiva y el muestreo representativo.

Recopilación exhaustiva

La recopilación exhaustiva se utiliza para adquirir todo el material producido en relación con un área temática, un periodo de tiempo o una región geográfica determinados. Este planteamiento requiere importantes recursos institucionales, o un enfoque limitado. El depósito legal de las publicaciones es quizás el planteamiento exhaustivo más conocido, que se da cuando una biblioteca nacional trata de compilar toda la producción editorial de la nación mediante un requisito legal para que los editores depositen ejemplares de cada publicación que producen. Los museos también pueden tratar de reunir todas las obras creadas en un periodo determinado.

Los archivos emplean prácticas de recopilación exhaustivas en casos concretos. Por ejemplo, pueden intentar recopilar todo lo relacionado con un personaje público influyente o un acontecimiento concreto de importancia nacional.

Muestreo representativo

El muestreo es otro método utilizado para identificar el material con miras a su conservación a largo plazo. Suele utilizarse cuando una institución no cuenta con los recursos o la capacidad para realizar una recopilación exhaustiva y cuando resulta problemático elegir el material basándose en criterios de selección específicos. En estos casos, el muestreo ofrece un medio para capturar una imagen representativa, haciendo que la selección y la conservación sean más fáciles de gestionar y requieran menos recursos.

Por ejemplo, una biblioteca nacional puede realizar rastreos programados de todo un dominio web nacional para preservar un retrato representativo de su presencia nacional en línea en diferentes momentos. Un archivo puede seleccionar los expedientes gubernamentales utilizando una metodología de muestreo, como conservar únicamente los casos más documentados, o los de un año determinado, o los que empiezan por una letra determinada del alfabeto, o los que determine un análisis estadístico.³

Considerar el contexto

Antes de emprender un proyecto de selección del patrimonio digital, puede ser útil analizar el entorno contextual general en el que se desenvuelve. Ante la creciente cantidad de información a la que se enfrentan los profesionales del patrimonio en un mundo digital, tener en cuenta el contexto a la hora de investigar y tomar decisiones sobre la selección puede ahorrar tiempo y recursos valiosos y puede aportar claridad a la adopción de decisiones, que corre el riesgo de enturbiarse en entornos digitales complejos.

Examinar el contexto de creación del patrimonio digital y las interrelaciones entre los creadores, las actividades y los sistemas puede facilitar la toma de decisiones a un nivel más

³ Algunos archivos gubernamentales han reconocido las limitaciones de este planteamiento. Véase Cook, 1991.

estratégico. Tener en cuenta el contexto mientras se navega por el árbol de decisión que se muestra a continuación puede resultar útil como principio rector para disminuir las tangentes y asegurarse de que las instituciones adquieran el patrimonio digital con los metadatos y la información sobre su procedencia esenciales para preservarlo y hacerlo accesible.

Con ese fin, se han incluido consideraciones contextuales ("comprobaciones de contexto") en el árbol de decisión para facilitar el análisis.

Árbol de decisión para la selección en una institución concreta

Este planteamiento consta de cuatro pasos, planteados en una serie de preguntas, para dar apoyo a una evaluación coherente y basada en pruebas:

Paso 1: Identificación

Defina los parámetros del proyecto, si procede. Determine si es suficiente una decisión del tipo *sí/no*, o si se requiere una evaluación relativa (alta, media, baja) para comparar los materiales.

Identifique el material que se quiere adquirir o evaluar.

- ¿Cuál es su título, creador, procedencia, extensión, condición?
- ¿Tiene la institución la capacidad y los derechos para capturar, conservar y poner a disposición todos los datos/ficheros/información?
- ¿Qué tipo de metadatos y en qué cantidad están disponibles? Si los metadatos se corrompen, ¿cómo se solucionará el problema?⁴

Documente su proceso de adopción de decisiones en este paso y conserve el registro. Asegure el mantenimiento constante del registro, así como su accesibilidad.

Comprobación del contexto: ¿se puede realizar la identificación a un nivel superior al del propio objeto digital? ¿Existen funciones o actividades que puedan orientar las decisiones sobre la selección?

Por ejemplo, los archivos gubernamentales pueden decidir no adquirir ningún registro creado por las actividades administrativas porque están suficientemente documentadas en el contexto de los instrumentos políticos de alto nivel y de documentos que apoyan la toma de decisiones en general.

⁴ Los metadatos pueden estar incompletos o corrompidos debido a la migración, pero esto no debería impedir que un registro sea seleccionado. Puede ser que la institución conserve el registro, pero complemente los metadatos o encuentre un mecanismo alternativo para ponerlo a disposición mientras encuentra una explicación para justificar los metadatos incompletos.

Paso 2: Marco jurídico

Comience por plantear las siguientes preguntas para identificar las cuestiones jurídicas relacionadas con la conservación a largo plazo de su material:

- ¿Tiene la institución la obligación jurídica de conservar el material? ¿Requieren el mandato o las políticas de la institución en materia de preservación digital y desarrollo de colecciones la conservación de este material?
- ¿Ha considerado los posibles obstáculos a la conservación, como los derechos de propiedad intelectual o las restricciones a la privacidad?

En caso afirmativo, consérvelo. Si se confirma una decisión positiva y no es necesaria ninguna otra selección, no se requieren más pasos.

Documente su proceso de adopción de decisiones en este paso y conserve el registro. Asegure el mantenimiento constante del registro, así como su accesibilidad.

Paso 3: Aplicación de los criterios de selección

Si se requiere una nueva selección, la institución puede evaluar el material utilizando tres criterios de selección: **importancia**, **sostenibilidad** y **disponibilidad**. Estos criterios deben evaluarse en el orden que resulte más eficiente o eficaz para la institución, comenzando generalmente por los criterios más fáciles de evaluar y siguiendo hasta llegar a una decisión final.

3 a): Importancia

Para determinar la importancia de un objeto, tendrá que plantearse una serie de preguntas sobre su autenticidad, origen, motivo de creación y valor para quienes lo han creado. Comience por lo siguiente:

- ¿En qué medida apoyan la misión y el mandato institucionales y se alinean con ellos?
- ¿El objeto digital nace en formato digital o sustituye a un objeto físico? ¿Sigue existiendo el objeto físico?
- ¿Tiene un valor social, cultural, histórico o artístico significativo para la comunidad a la que sirve la institución, para la comunidad que creó el objeto o para la humanidad en su conjunto?
- ¿Tiene un valor informativo, de contenido, de uso o de investigación significativo?
- ¿Cómo se verán afectadas las partes interesadas de la institución (clientes, patrocinadores, sociedad) si no se conserva este objeto digital?
- ¿Ha sido creado el material del patrimonio documental por indígenas o por una comunidad infrarrepresentada, o se refiere a ellos? De ser así, las consultas con dichas comunidades deben ser consideradas parte del proceso de selección.

Si el objeto digital es importante según estos criterios, considere la posibilidad de conservarlo. Documente su proceso de adopción de decisiones en este paso y conserve el registro. Asegure el mantenimiento constante del registro, así como su accesibilidad.

Comprobación del contexto: ¿Qué información existe para identificar actividades, funciones o a creadores que puedan ayudar con la selección?

Por ejemplo, ¿existen análisis empresariales o historiales administrativos que puedan aportar información sobre por qué y cómo se creó el material del patrimonio digital? ¿Puede esto ayudar a identificar lugares o puntos de creación que pueda ser crucial documentar? ¿Cuáles son las relaciones entre las distintas actividades, funciones o creadores que pueden revelar áreas de valor importantes?

3 b): Sostenibilidad

Antes de seleccionar el patrimonio digital para su conservación a largo plazo, debe analizar su propia capacidad para proporcionar un acceso sostenible al material.

Las preguntas que se proponen aquí son referencias para reflexionar. Si una institución no tiene los recursos, las capacidades, la capacidad técnica o el presupuesto para conservar y hacer accesible el material, esto no significa necesariamente que no deba adquirir ese material. Puede que tenga que idear soluciones provisionales a corto plazo para permitir la conservación y/o el acceso. Estas decisiones deberán adoptarse caso por caso, a discreción de la institución y de las partes interesadas que participen en el proceso.

Es importante recordar que la preservación digital no es solo un problema técnico que requiere únicamente soluciones técnicas. Los conservadores, los gestores de las colecciones, los administradores de la información y los demás responsables de la preservación de las colecciones físicas deben comprender las necesidades de preservación del material digital. Garantizar que todas las jurisdicciones comprendan los objetivos de recopilación de las organizaciones patrimoniales y la ética que sustenta el cuidado de sus colecciones digitales ayudará a hacer frente al desafío cada vez mayor que supone construir y mantener estas colecciones en el tiempo.

En todos los casos, la evaluación de la "sostenibilidad" debe estar basada en consideraciones relativas a la "importancia" y el "contexto". Empezar por plantearse las preguntas siguientes:

- ¿Dispone la institución de infraestructura suficiente para preservar los materiales del patrimonio digital? De no ser así, ¿existe un plan para imprevistos que garantice la sostenibilidad del material?
- ¿Tiene la institución capacidad técnica para abarcar, migrar y conservar el patrimonio digital?
- ¿Se requieren derechos específicos para transferir o migrar el material a diferentes formatos de archivo y/o soportes físicos?
- ¿Se dispone de suficientes metadatos para acceder al patrimonio digital y preservarlo?

- ¿ Es adecuado el formato del objeto digital para su conservación o debe ser migrado?⁵
- ¿Puede la institución hacer que el material sea accesible para la investigación, la exposición o cualquier otro uso que satisfaga las expectativas del público?

Documente su proceso de adopción de decisiones en este paso y conserve el registro. Asegure el mantenimiento constante del registro, así como su accesibilidad.

Comprobación del contexto: ¿Cómo se ha creado y utilizado el patrimonio digital a lo largo del tiempo? ¿Existen vínculos entre sistemas o conjuntos de datos que puedan alterar la valoración del patrimonio digital o crear problemas de conservación?

Comprender el control intelectual del material a lo largo de su vida puede proporcionar pistas valiosas para determinar si es adecuado para su adquisición. Por ejemplo, comprender los sistemas que utilizaba una organización y las estructuras de clasificación que se empleaban puede ser esencial para entender por qué el material del patrimonio digital tiene el aspecto que tiene.

Un sistema puede haber sido adaptado para su utilización de una forma distinta al uso para el que fue concebido. Pensar en el contexto de creación puede permitir a una institución seleccionar el material del patrimonio digital lo antes posible en su ciclo vital, idealmente antes de su creación.

Esto puede crear oportunidades para que las instituciones de la memoria trabajen con los creadores, prestando asesoramiento sobre la manera de establecer su infraestructura y formatos desde el principio para apoyar la conservación a largo plazo.

Redundancia

Deben existir múltiples copias del patrimonio digital importante, incluidos los ficheros maestros con los metadatos correspondientes, almacenadas en al menos dos ubicaciones físicas diferentes. Las instituciones culturales pueden utilizar una combinación de almacenamiento *in situ*, externo y basado en la nube, pero los originales digitales deben tener una copia de seguridad en al menos otra ubicación. Los lugares de almacenamiento deben elegirse de manera que se reduzca el riesgo de pérdidas debidas a catástrofes naturales o provocadas por el hombre y a crisis económicas o políticas.

Gestión activa

Las instituciones culturales deben gestionar activamente sus activos patrimoniales digitales para garantizar su accesibilidad e integridad a largo plazo. El patrimonio digital debe conservarse en formatos de archivo abiertos y bien documentados, sin encriptación y, al menos, de compresión sin pérdidas. Este es el método que se recomienda encarecidamente a las instituciones culturales para la gestión activa de los objetos digitales. El almacenamiento debe utilizar dos o más tipos diferentes de medios de almacenaje, desde servidores

⁵ Aunque hay argumentos a favor de conservar todos los formatos, esta decisión dependerá de los recursos y la infraestructura del profesional encargado de tomar la decisión.

institucionales hasta medios portátiles (por ejemplo, discos magnéticos, medios ópticos o cintas magnéticas).

Los fallos de los sistemas a largo plazo pueden provocar la pérdida de información vital del patrimonio digital almacenado. Muchas instituciones se protegen contra esto realizando una actualización periódica de los medios, que consiste en leer los datos digitales, comprobar si hay errores utilizando técnicas de corrección de errores y reescribirlos en nuevos medios. Para evitar fallos en el *software*, los propietarios de datos digitales suelen utilizar protocolos basados en normas para el acceso al almacén de datos, que consisten en la ejecución de distintas aplicaciones del *software* de almacenamiento en los diferentes sitios de almacenamiento. Por lo tanto, la integridad y la fiabilidad de los datos no dependen de la integridad y la fiabilidad de una sola aplicación.

3 c): Disponibilidad

Considere la posibilidad de que el patrimonio digital esté disponible en otras instituciones de la comunidad o red del patrimonio. Comience por plantearse las siguientes preguntas:

- ¿Es esta institución la única que conserva este material, o hay copias exactas en otras instituciones? ¿Sirve de referencia a los fondos?⁶ ¿Sirve de referencia al conjunto de la colección institucional? ¿Es inusual o único, o ha sido ampliamente duplicado?
- ¿Dónde se utilizará más o será más beneficioso para el público?
- ¿Está expuesto a algún riesgo en otras instituciones?
- ¿Es esta institución la más adecuada o la que reúne mejores condiciones para conservar y hacer accesible este patrimonio digital?
- Tenga en cuenta que es necesario un cierto grado de redundancia para garantizar la correcta conservación del patrimonio digital. Puede consultar información adicional sobre la redundancia en el Apéndice 6.

La disponibilidad, al igual que la sostenibilidad, debe basarse en la importancia y el contexto, en particular si los ficheros que se están analizando se refieren a una comunidad infrarrepresentada o han sido creados por ella.

Documente su proceso de adopción de decisiones en este paso y conserve el registro. Asegure el mantenimiento constante del registro, así como su accesibilidad.

Comprobación del contexto: ¿Existen relaciones importantes entre las actividades, las funciones o los creadores que puedan ayudar en la selección?

Paso 4: Decisión

Recopile y revise toda la documentación generada durante el proceso y tome una decisión basada en los resultados de los pasos 1 a 3.

Este enfoque es flexible; no todas las preguntas serán aplicables a cada institución. El orden de los criterios tampoco es inamovible; por ejemplo, en algunos casos, el paso 3 c) podría

⁶ Término archivístico que hace referencia a los registros agregados de un determinado creador.

evaluarse mejor antes de los pasos 3 a) y 3 b), en particular si otra institución es a todas luces más adecuada.

Documente y registre el fundamento y la justificación de la evaluación o decisión. Esto es vital, tanto para la gobernanza como para capturar información importante que puede resultar útil en una posible reevaluación en el futuro. Prepare una declaración escrita sobre la importancia del patrimonio digital, el contexto de creación y sus problemas técnicos de preservación, incorporando las respuestas a las preguntas de los pasos 1 a 3. Los argumentos que sustentan la decisión son a menudo más importantes que la propia evaluación. Debería crearse un formulario de evaluación institucional estándar o un documento de valoración para recoger estos argumentos y dejar constancia de la decisión.

A pesar de estas posibles variaciones en su aplicación, este planteamiento debería ayudar a las instituciones culturales a adoptar mejores decisiones a la hora de seleccionar el material digital para su conservación a largo plazo.

Parte 2: Análisis detallado de la recopilación en el entorno digital

A la hora de establecer un programa de selección y preservación digital, puede resultar útil preguntarse cómo puede afectar la recopilación de contenidos digitales a los procesos internos de la institución. También resulta útil examinar otras cuestiones externas, como las dificultades legales y el papel de las asociaciones, que son dos aspectos que repercuten en la preservación digital a largo plazo.

Selección digital en bibliotecas, museos y archivos

En el mundo digital, los modelos de operación y las modalidades para servir al público que utilizan las bibliotecas, los museos y los archivos están cambiando rápidamente. En este contexto digital, es importante que los profesionales de la información no dejen sus actividades tradicionales de gestión de colecciones en manos de los tecnólogos.

El ámbito digital supone un cambio de paradigma y un reto innegable para el mandato de recopilación y preservación, que requiere un replanteamiento de la forma en que las instituciones patrimoniales identifican la importancia y evalúan el valor. Esto también pone de relieve la necesidad de que los profesionales de las bibliotecas, los archivos y los museos consideren la recopilación digital junto con los procesos tradicionales de gestión de colecciones, reconociendo al mismo tiempo que cada uno de ellos presenta características y planteamientos únicos.

Si bien algunas de las fronteras entre las bibliotecas, los archivos y los museos se están desdibujando en la era digital, sigue habiendo cuestiones específicas que hay que tener en cuenta y que tienen una pertinencia específica para cada una de estas instituciones.

Bibliotecas

Las bibliotecas se enfrentarán al reto de la selección digital con respecto a las publicaciones electrónicas, la captura de sitios web y los contenidos privados en los sitios de los medios sociales como Facebook y YouTube. Las bibliotecas nacionales que aspiran a crear una colección completa, a menudo con una fuerte tradición de depósito legal, tendrán que realizar la selección de publicaciones más efímeras en formato digital. En el pasado, la selección la hacían, en efecto, los editores que "conservaban" la producción creativa mediante selecciones editoriales que determinaban lo que se publicaría. En el mundo democratizado de la autoedición y los libros electrónicos, las bibliotecas nacionales tendrán que modificar los planteamientos integrales del pasado y adoptar criterios de selección para la conservación a largo plazo.

Aunque no todas las bibliotecas poseen colecciones de patrimonio cultural reconocidas como importantes en el marco nacional o internacional, muchas de ellas son depositarias de la memoria local o regional. Es cada vez más probable que el registro de esta memoria, como los periódicos y las grabaciones audiovisuales, sea un material de origen digital.

Habida cuenta de que las bibliotecas tienen el mandato de satisfacer las necesidades de su comunidad, la accesibilidad y el uso a largo plazo son esenciales para su preservación. Las colecciones digitales deben establecerse teniendo en cuenta los principios básicos de la funcionalidad y los medios de descubrimiento. Para algunas bibliotecas, la selección de objetos digitales para su conservación a largo plazo puede centrarse principalmente en la evaluación de las publicaciones que ya forman parte de su colección, adquiridas originalmente para su uso a corto plazo, en lugar de evaluar nuevas publicaciones para su adquisición.

Independientemente de su tamaño, la mayoría de las bibliotecas se beneficiarán de la codificación de su proceso de selección digital y de la estructuración de sus criterios de selección atendiendo a valores similares compartidos entre las instituciones de la memoria. Esto ayudará a las bibliotecas a identificar los tipos de material digital que pueden formar parte de su mandato a fin de seleccionarlos para su conservación a largo plazo, evitando tanto la duplicación como la posible pérdida de material.

Museos

Los museos que cuentan con colecciones importantes y bien desarrolladas de material cultural físico suelen realizar adquisiciones para su preservación permanente y adoptan decisiones de desarrollo de las colecciones en este contexto. Esta cultura material se está volviendo cada vez más digital: por ejemplo, las máquinas que funcionan con programas informáticos, las obras de arte de origen digital, las instalaciones artísticas digitales y la documentación digital de yacimientos arqueológicos. La información de las investigaciones relacionadas con los fondos físicos de los museos también está cada vez más digitalizada.

Además, los museos están recopilando patrimonio digital con un importante componente físico, por ejemplo, teléfonos móviles, tabletas, ordenadores y videoconsolas. Estas decisiones relativas a las colecciones de los museos pueden basarse en el enunciado de su

misión. Por ejemplo, un museo de diseño puede albergar un "icono de estilo" que es en parte físico y en parte digital, o un museo de ciencias albergar *hardware* importante (y *software* relacionado).

Por lo tanto, el patrimonio digital en los museos puede dividirse en las categorías siguientes: objetos de origen digital en la colección, información digital o digitalizada sobre la colección y representaciones digitales de los objetos físicos de la colección (imágenes digitales o escaneos en 3D, por ejemplo). En virtud de esta categorización, los museos deberían dar prioridad a las categorías primera y segunda para su preservación a largo plazo. Las categorías segunda y tercera también incluyen la documentación administrativa generada por la institución.

No se puede destacar lo suficiente la importancia que tienen los metadatos (información sobre el patrimonio físico y digital) para los museos. Los metadatos incluyen información contextual creada sobre el patrimonio físico y digital antes de su entrada en el museo e información contextual creada durante su vida en el museo. El principio de procedencia también es importante para los museos.

Archivos

Los archivos se centran en la importancia de la autenticidad, la procedencia y el contexto a la hora de valorar los ficheros para su adquisición. El entorno jurídico suele dictar qué información digital debe adquirir un archivo y cómo puede ponerse a disposición del público y la investigación, o si esto es posible. En algunos casos, el material de origen digital con licencia abierta puede archivar de forma masiva, dando lugar a funciones nuevas o evolucionadas para los propios archivos.

Los archivos adquieren documentos originales o únicos para su preservación permanente y, tradicionalmente, se han basado en el tiempo transcurrido entre su creación (por ejemplo, 20 años) y su adquisición para aportar una perspectiva histórica a la hora de adoptar decisiones de selección. Sin embargo, la rápida obsolescencia de los formatos digitales y la incapacidad de los creadores de los ficheros y los datos para mantener la información, los medios de almacenamiento y los sistemas de *hardware* y *software*, están mermando las opciones de selección. Cada vez son más los archivos que empiezan a enfrentarse a la necesidad de recopilar con mayor antelación ficheros y datos de origen digital, medios sociales, software para la web, inteligencia artificial y muchas otras formas de datos importantes, lo que requiere una interacción más temprana con los creadores de los registros para influir en el diseño y la producción de ficheros y datos de interés a largo plazo. Esto también ha obligado a establecer relaciones de trabajo más estrechas con los tecnólogos de la información, los científicos de datos y muchos otros agentes para garantizar la integridad del proceso de creación de registros y datos y el mantenimiento de la información hasta que pueda transferirse. La intervención temprana de la comunidad es necesaria porque la forma en que se crean los registros y los datos influirá en su adquisición, conservación y acceso.

La capacidad de los archivos varía considerablemente, pero esto no significa que deban abstenerse de conservar los materiales digitales (ficheros, datos, etc.) si carecen de presupuesto, capacidad e infraestructura. Existen muchos modelos de madurez que ayudan

a los archivos no solo a evaluar sus capacidades, sino también a identificar una hoja de ruta práctica que tenga en cuenta los recursos de que disponen.⁷

Estos factores que influyen en el entorno de selección no son necesariamente exclusivos de cada una de las comunidades de bibliotecas, museos y archivos. De hecho, es más que probable que se produzca algún solapamiento. Sin embargo, el análisis de la diversidad de nuestras comunidades ayuda a esclarecer la variedad de problemas a los que deben enfrentarse las instituciones a la hora de identificar y seleccionar el patrimonio para su conservación a largo plazo.

Alianzas para la selección y preservación digital

Las instituciones patrimoniales y de la memoria pueden beneficiarse del intercambio de conocimientos y de la coordinación de esfuerzos en materia de recopilación y preservación digital mediante la participación en redes nacionales e internacionales.

Cooperación internacional

La participación en intercambios internacionales conjuntos y de experiencias sirven para coordinar los esfuerzos de recopilación y preservación digital. Algunos países ya habrán establecido estrategias y planes nacionales que pueden servir de modelo en otros lugares. Las instituciones con experiencia en la gestión de colecciones digitales pueden compartir sus conocimientos sobre cuestiones como la legislación, las normas y las prácticas, los sistemas de selección y acceso digital, la conservación del arte digital, la infraestructura de preservación y la seguridad.

Los archivos, museos y bibliotecas colaboran estrechamente con sus propios consejos y federaciones internacionales para la cooperación técnica y los intercambios. Se debe consultar con organizaciones internacionales de la sociedad civil, como el ICA, la IFLA y el ICOM, y con organizaciones intergubernamentales, como la UNESCO, ya que la preservación del patrimonio digital es un reto mundial, y estas organizaciones pueden liderar las conversaciones en el ámbito internacional. También son plataformas sólidas en las que desarrollar estrategias conjuntas.

El papel de las instituciones nacionales

Las instituciones nacionales de los sectores de las bibliotecas, los archivos, los museos y patrimonial deben desempeñar un papel fundamental de liderazgo de los organismos gubernamentales y las comunidades del patrimonio en cuestiones de selección y preservación digital.

En muchos países, algunas instituciones nacionales designadas cuentan con legislación y políticas relativas al desarrollo de colecciones y la gestión del patrimonio cultural. La función

⁷ Algunos ejemplos de modelos de madurez de la preservación digital son los niveles de preservación digital de la NDSA (<https://ndsa.org/publications/levels-of-digital-preservation/>) y el modelo de evaluación rápida del CPD (<https://www.dpconline.org/digipres/dpc-ram>).

de depósito legal de las bibliotecas nacionales permite compilar las obras publicadas, y los archivos nacionales tienen el mandato de adquirir los documentos oficiales de sus gobiernos. Los museos también han empezado a ocuparse de la recopilación de objetos digitales en el marco de sus políticas de coleccionismo contemporáneo.

La escala y los desafíos únicos que plantean los contenidos digitales hacen que las instituciones nacionales asuman de forma natural un papel de liderazgo en la gestión de contenidos digitales. Estas instituciones cuentan con normas, infraestructura y sistemas que pueden ampliarse para apoyar al resto de la comunidad patrimonial, en particular a las bibliotecas, galerías y archivos más pequeños.

Las instituciones nacionales también deben ampliar su red con objeto de reunir los conocimientos especializados necesarios para elaborar un plan coordinado y de colaboración destinado a desarrollar estrategias y normas nacionales de selección, recopilación y preservación. En la red pueden participar creadores de contenidos, expertos de la industria, coleccionistas de patrimonio y organismos gubernamentales clave que se ocupan de las innovaciones tecnológicas, la protección de datos, los derechos de autor y los medios digitales.

A nivel de las instituciones nacionales, es importante que las bibliotecas, los archivos, los museos y otras instituciones que se ocupan del patrimonio se pongan de acuerdo para definir las funciones y el alcance de las colecciones nacionales, incluido un plan cooperativo de recopilación y conservación. Es recomendable que la selección de los contenidos del patrimonio digital y su gestión a largo plazo sean un esfuerzo de colaboración nacional, habida cuenta de los retos y la escala de los contenidos digitales que se crean en un país, que quizás ya no puedan ser gestionados por una única institución. Por ejemplo, si una institución no cuenta con los recursos, las capacidades o la capacidad técnica o presupuestaria necesarios para preservar algún material, podría trabajar en colaboración con otras instituciones con más capacidad a nivel nacional o regional que puedan ayudar a facilitar las recomendaciones.

La importancia de la promoción

En el momento de redactar este documento, las restricciones en vigor debidas a la COVID-19 plantean un reto determinante, pero los recortes presupuestarios generales y otras dificultades seguramente seguirán siendo motivo de preocupación. Habida cuenta de los numerosos retos a los que se enfrenta el sector público, es más importante que nunca concienciar al público y a los responsables de la toma de decisiones sobre el valor cultural de las colecciones de patrimonio digital.

Estas actividades de promoción deben tener como objetivo insistir en los beneficios derivados del acceso a largo plazo al material digital. Por ejemplo, se puede destacar el mandato legal de hacer accesible al público el patrimonio cultural, o hacer hincapié en los beneficios económicos, intelectuales y científicos que conlleva el acceso a estos contenidos digitales.

Promoción a nivel institucional

A los niveles local o nacional, la promoción general puede ser relativamente neutra y adoptar la forma de relaciones públicas, campañas publicitarias, medidas de sensibilización y seminarios profesionales.

Tal vez la mejor promoción que puede hacer una institución de la memoria sea la celebración de actos públicos informativos, con personal bien informado y motivador. A largo plazo, esta medida puede servir para obtener el apoyo del público, desmitificar la institución, disipar malentendidos y comunicar cómo esta labor beneficia a la sociedad en general.

Alianzas con la sociedad civil

La cooperación entre los profesionales y las organizaciones de la sociedad civil puede ser eficaz en este sentido, ya que no siempre es conveniente que las instituciones nacionales asuman un papel de liderazgo en los esfuerzos de promoción y cabildeo. Organizaciones internacionales como la IFLA, el ICA y el ICOM, y organizaciones intergubernamentales como la UNESCO participan activamente en campañas de promoción relacionadas con la preservación del patrimonio y el acceso al mismo que incluyen el formato digital.

El éxito de las actividades de promoción puede influir en la dirección de los debates intergubernamentales, puede dar lugar a normas técnicas eficaces o puede servir para obtener apoyo y financiación para realizar actividades profesionales.

La influencia del entorno legal

El entorno legal tiene gran influencia en la selección, la preservación, y la disponibilidad y accesibilidad del patrimonio digital. Las leyes internacionales y nacionales varían mucho, pero normalmente regulan la captura, difusión, duplicación, acceso y uso del patrimonio digital. Sin embargo, internet trasciende las fronteras territoriales, por lo que a menudo resulta difícil determinar qué legislación aplicar e identificar a los titulares de los derechos.

El **depósito legal** en las bibliotecas permite capturar, preservar y hacer accesible el material digital publicado (véase la lista de comprobación de la IFLA sobre el depósito legal para obtener más información al respecto). Por su parte, los archivos se rigen por la **legislación archivística** o por las **políticas de adquisición**.⁸ La legislación archivística encomienda a los archivos el mandato de recopilar y gestionar los documentos públicos digitales procedentes de los ministerios o departamentos y dependencias gubernamentales. En las organizaciones del sector privado, es más frecuente que la recopilación, la preservación y la accesibilidad de los materiales digitales se rijan por una política de adquisiciones. Por el contrario, las colecciones de un museo pueden no estar reguladas por la legislación. Los museos nacionales suelen trabajar con mayor frecuencia dentro de un marco legal más definido. Sin embargo, la situación varía de un país a otro o puede variar según la gobernanza del museo.

⁸ <https://www.ifla.org/publications/node/93470>.

Un marco legislativo sólido y favorable es una condición esencial para el éxito de un programa de selección y preservación digital. En sus Directrices para la Conservación del Patrimonio Digital de 2003, la UNESCO señalaba que:

"Hacer que la legislación sobre archivos, así como el depósito legal o voluntario en bibliotecas, archivos, museos u otras instituciones públicas de conservación, se aplique al patrimonio digital, ha de ser un elemento esencial de la política nacional de preservación. La legislación sobre el derecho de autor y derechos conexos debería permitir a dichas instituciones llevar a cabo el proceso de conservación en un marco legal".

Por lo tanto, es importante que las organizaciones que pongan en marcha un programa de preservación digital se aseguren de que el contexto legislativo en el que operan les permita capturar los contenidos digitales para su conservación.

Hay que tener en cuenta

¿Cubren las leyes de depósito legal y la legislación sobre archivos los contenidos digitales? En caso afirmativo, ¿de qué tipo? ¿Impiden las leyes actuales la adquisición de patrimonio documental de importancia nacional? En caso afirmativo, ¿cómo pueden garantizar las instituciones la selección y la conservación?

¿Restringen las leyes de depósito legal la recopilación de obras accesibles al público, o sería posible recopilar o solicitar obras que son de pago o tienen otro tipo de restricciones?

¿Qué normas existen sobre el tratamiento de las obras que posteriormente no están disponibles o se restringen en línea?

Otras consideraciones legales: derechos de autor y privacidad

La legislación sobre derechos de autor, salvo excepciones y limitaciones específicas, puede prohibir la realización de copias. Esto plantea nuevos problemas en un entorno digital en el que puede ser necesaria la duplicación para la conservación a largo plazo. Además, los materiales digitales suelen depender de un *software* para su búsqueda y recuperación, y ese *software* también puede estar protegido por derechos de autor.

Algunos países han promulgado leyes que prohíben eludir las medidas tecnológicas de protección utilizadas para impedir la copia y redistribución, lo que dificulta la conservación e impide el futuro acceso legítimo al patrimonio digital. Otros países tienen disposiciones jurídicas que les permiten eludir las medidas tecnológicas de protección con fines de conservación.

Hay jurisdicciones que cuentan con **disposiciones de "uso leal"**, que algunas organizaciones utilizan para recopilar material que se considera de dominio público (por ejemplo, sitios web de acceso público). Esto significa que este material puede ponerse a disposición de los usuarios sin necesidad de solicitar una autorización previa.

Otras jurisdicciones utilizan una "**renuncia**" de la misma manera, y es necesario que el propietario de los derechos de autor emita una objeción antes de que el material esté disponible.

La comprensión de las implicaciones legales pertinentes de la preservación digital en su jurisdicción local hace que los profesionales estén mejor preparados para seleccionar el patrimonio digital con fines de preservación a largo plazo y elaborar un marco de responsabilidad legal a tal efecto.

Las leyes de privacidad también pueden afectar a la disponibilidad y accesibilidad de la información digital que contiene datos personales. Esto dependerá de la jurisdicción, pero las leyes sobre el "derecho al olvido", por ejemplo, pueden afectar a lo que las organizaciones que se ocupan del patrimonio cultural pueden adquirir, conservar y poner a disposición.

La falta de coherencia internacional en lo que respecta a las leyes de derechos de autor, y la falta de claridad sobre las modalidades de cooperación a través de las fronteras pueden tener repercusiones no deseadas en los esfuerzos por preservar las obras (por ejemplo, a través de redes de preservación digital), almacenarlas (por ejemplo, en la nube) y dar acceso a las mismas, todo lo cual afecta a las decisiones de selección de contenidos.

Para apoyar la preservación de los contenidos digitales, un objetivo clave a largo plazo es una acción internacional que permita salvaguardar los materiales y dar acceso a los profesionales del patrimonio, junto con los avances a nivel nacional para garantizar que las decisiones de los gestores de colecciones digitales se rijan únicamente, en la medida de lo posible, por criterios profesionales.

Hay que tener en cuenta: ¿Autoriza claramente la legislación sobre el depósito legal a las instituciones a preservar el material mediante copia, en la forma que sea necesaria?

¿Pueden ignorar las instituciones las cláusulas contractuales y eludir las medidas de protección tecnológica?

¿Cuáles son las normas para dar acceso a las obras digitales recopiladas?

¿Cuáles son las disposiciones de preservación de los datos personales? ¿Existen leyes de privacidad o protección de datos? ¿Cómo influyen en lo que se puede adquirir, conservar y poner a disposición?

Conclusión

Además de abrir un mundo de posibilidades para la expresión y el ingenio humanos, la tecnología digital también presenta el reto único de seleccionar lo que seguirá siendo accesible. Del mismo modo, a medida que evoluciona la tecnología utilizada para crear estas expresiones, también deben hacerlo los planteamientos de selección, junto con las metodologías de conservación y acceso.

Las intervenciones y decisiones de los profesionales de hoy influirán en la memoria para el futuro. Se trata de una enorme responsabilidad que las instituciones que se ocupan de la memoria y el patrimonio afrontan en colaboración con otras instituciones y legisladores de ámbito nacional, así como con redes internacionales y organizaciones de la sociedad civil.

Deben establecerse estrategias para seleccionar este material digital con miras a su conservación a largo plazo, codificar el proceso de selección y recopilación y documentar la metodología de adopción de decisiones. Se invita a los profesionales a basarse en los criterios de selección y los antecedentes que aquí se ofrecen y adaptarlos a su contexto. En los apéndices que figuran a continuación se profundiza en algunos conceptos clave que pueden aportar más información a la hora de crear una estrategia de selección.

Las Directrices se han redactado teniendo en cuenta el panorama en constante cambio de la creación de información que define nuestra época. Considere estas Directrices como parte de un esfuerzo internacional más amplio para preservar y facilitar el acceso al patrimonio digital mundial, un esfuerzo que seguramente seguirá evolucionando en el futuro.

Apéndice 1: Coleccionismo digital frente a coleccionismo tradicional

En este capítulo se analizan con mayor detalle las similitudes y diferencias entre el coleccionismo tradicional y el digital. Se incluye un ejemplo de gestión de riesgos para demostrar cómo afectan estas diferencias a la preservación en términos prácticos.

Similitudes

La recopilación digital no es "un problema técnico que requiere únicamente soluciones técnicas. Se trata más bien de un problema social, cultural y organizativo, al igual que los enfoques conceptuales y éticos tradicionales de las [colecciones] físicas se derivaron de las preocupaciones sociales y culturales" (Slade, S., Pearson, D., y Knight, S., 2019).

Al igual que el coleccionismo físico, el coleccionismo digital "se caracteriza por la procedencia, el contexto, la cadena de custodia y la documentación exhaustiva a lo largo del tiempo. El entorno de preservación digital puede considerarse del mismo modo que el entorno de almacenamiento de las colecciones físicas. El almacenamiento para la preservación de las colecciones digitales no consta únicamente de la sala y el equipo de almacenamiento, sino también de la organización de los objetos de la colección, los métodos para localizarlos con precisión, la seguridad que se les proporciona, las políticas sobre su preservación y cuidado, el entorno en el que se ubican y los métodos de supervisión del entorno para tener bajo control cualquier riesgo para su preservación" (Slade, S., Pearson, D., & Knight, S., 2019).

Los riesgos para las colecciones tanto físicas como digitales son similares, pero difieren en la forma en que se manifiestan (véase el cuadro 1).

Diferencias

No hay que subestimar el esfuerzo y la inversión necesarios para salvaguardar los materiales digitales. El esfuerzo será continuo e intensivo, y seguirá ampliándose a medida que se complican la escala, el formato y los problemas de accesibilidad.

Los problemas concretos del coleccionismo digital incluyen:

- Formatos obsoletos (por ejemplo, WordStar)
- Material digital en soportes físicos no compatibles (por ejemplo, disquetes o Betacam)
- Modelos de publicación exclusiva en la nube
- Jurisdicciones de plataformas modernas: material digital con origen en plataformas extranjeras no sujetas a la legislación local de recopilación (por ejemplo, Facebook, YouTube, Instagram o Twitter)
- Falta de capacidad o espacio en el sistema de conocimientos
- Modelos de gestión de la propiedad intelectual obsoletos (por ejemplo, derechos de autor) que requieren nuevos modelos de acceso, concesión de licencias y uso o reutilización
- Falta de una respuesta coherente a los problemas a largo plazo de la protección de los activos digitales para la innovación, el uso y la reutilización, y la reproducibilidad (es decir, poder validar la integridad de los resultados de la investigación)
- La escasa gestión de la información digital en el sector público limita las oportunidades para aprovechar al máximo el valor de la información gubernamental.

- Material digital que está compuesto por elementos independientes (ficheros), o los necesita, que no están necesariamente disponibles (legalmente) o archivados

Cuadro 1 - Ejemplo de gestión de riesgos de las similitudes entre el coleccionismo digital y el coleccionismo tradicional (Pearson, 2012).

| Riesgo | Definición | Ejemplo digital | Ejemplo físico |
|-------------------------------------|---|---|---|
| Riesgos basados en parámetros | Un criterio identificado por el personal para indicar un riesgo de preservación | Vídeo codificado con un códec problemático | La identificación de la película como nitrato de celulosa |
| Riesgo de excepciones | El valor de un parámetro supervisado queda fuera de un conjunto de valores aceptables | Un fichero en un formato determinado no supera la validación | Los niveles de humedad relativa dentro de una tienda no se ajustan a los niveles acordados |
| Riesgos de cambio | Un cambio en el estado de un parámetro supervisado para el contenido | La confianza en la identificación del formato de un archivo concreto ha cambiado, o una suma de comprobación ha fallado | El desvanecimiento de los colores en una fotografía en color o una acuarela debido a la exposición prolongada a la luz visible |
| Riesgos de conflicto | Uno o más instrumentos reflejan valores contradictorios para el parámetro | Una identificación de formato de fichero revela valores contradictorios | El sistema de gestión del edificio y un control ambiental independiente presentan diferentes lecturas de temperatura y humedad relativa para el mismo espacio de almacenamiento de la colección |
| Riesgos de valor desconocido | Valores indeterminados para los parámetros definidos | Hay valores desconocidos respecto del formato de fichero y la versión | Una silla tapizada muestra signos de deterioro; no se puede determinar la causa, que puede provenir del material utilizado durante la construcción de la silla, que ya no está disponible |
| Riesgos de compatibilidad de acceso | Cambios en el grado de compatibilidad que afectan a la capacidad de la organización para acceder a los contenidos de conformidad con su plan de preservación. | Una reducción significativa en la disponibilidad de <i>software</i> compatible para un formato de archivo concreto | Un agente tensioactivo utilizado para el lavado de textiles que ya no está disponible debido a los cambios en la legislación de las directrices medioambientales |
| Riesgos basados en el contenido | Características de los contenidos que pueden no ser identificables a partir de los metadatos | La presencia de etiquetas HTML obsoletas | Pintura anterior bajo la superficie de un óleo; no hay información al respecto en los archivos de procedencia de la pintura en esta colección y no se conoce su existencia |

Apéndice 2: Análisis detallado del código fuente del *software*

El *software* impregna nuestra vida personal y social. Encarna una gran parte del conocimiento tecnológico que impulsa nuestra industria, apoya la investigación moderna, cumple la función de mediador para acceder a los contenidos digitales y alimenta la innovación (UNESCO et al, 2019).

El *software* está escrito por humanos en el formato de *código fuente del software*. Se trata de una forma valiosa y única de conocimiento que, además de poder traducirse fácilmente a una forma ejecutable por una máquina, también está "escrita para que la lean los humanos" (H. Abelson et al, 1985) y "permite entrar en la mente del diseñador" (Shustek, 2006). A pesar de ser omnipresente, con frecuencia, el código fuente del *software* es considerado un simple accesorio de los programas informáticos "ejecutables", que se ejecutan en ordenadores personales o por conducto de servicios en la nube basados en Internet. Si bien el auge del *software* libre y de código abierto allanó el camino para el desarrollo de plataformas de alojamiento de código fuente accesibles al público, una parte importante del código fuente del *software* sigue siendo inaccesible, está en manos de empresas privadas, o simplemente no se comparte por diversos motivos.

Como se afirma en el llamamiento de París para que el código fuente de los programas informáticos se considere patrimonio para el desarrollo sostenible, es fundamental preservar a largo plazo este valioso patrimonio técnico, científico y cultural (UNESCO & INRIA, 2009). En otras palabras, el código fuente de los programas informáticos es hoy en día un artefacto digital peculiar que encarna el conocimiento y el esfuerzo que se dedica a la fabricación de herramientas y creaciones digitales que configuran el mundo digital actual, incluidos la investigación y el archivo. En un mundo digitalizado, el código fuente de los programas informáticos es un producto esencial de la investigación y debe considerarse, junto con las publicaciones y los datos de investigación, un pilar de la ciencia abierta.

Hasta hace muy poco, con la apertura del archivo Software Heritage (softwareheritage.org), el código fuente de los programas informáticos carecía tanto de un lugar para almacenar de forma permanente el código fuente como objeto digital, como de los mecanismos para identificar, almacenar y referenciar eficientemente los objetos digitales del código fuente, por ejemplo, mediante identificadores persistentes intrínsecos (Software Heritage, 2020).

Lecturas seleccionadas

Bussi, L., Di Cosmo, R., Montangero, C., Scatena G. (2019). *The Software Heritage Adquisition Process* UNESCO.

<https://unesdoc.unesco.org/ark:/48223/pf0000371017/PDF/371017eng.pdf.multi>

Di Cosmo, R. (2020). *How to use Software Heritage for archiving and referencing your source code: guidelines and walkthrough*.

<https://annex.softwareheritage.org/public/guidelines/archive-research-software.pdf>.

Apéndice 3: Análisis detallado de los datos de investigación

La investigación científica es el primer elemento que contribuye al progreso del conocimiento y, por ende, al progreso de la humanidad. Los países y las organizaciones internacionales de todo el mundo dedican una parte importante de su presupuesto anual a sostener, fomentar y mejorar las actividades de investigación científica. Las colaboraciones internacionales entre científicos son habituales, así como las iniciativas internacionales a gran escala en campos como la investigación espacial, la tecnología nuclear y la medicina. Todas estas actividades de investigación generan gran cantidad de datos en bruto que pueden necesitar varios años, si no decenios, para ser descifrados, compilados, comparados y analizados antes de llegar a cualquier conclusión científica, creando así nuevos conocimientos.

Los datos de investigación abiertos permiten que otros investigadores, además del creador de los datos, los reutilicen. Los usos de los datos de investigación abiertos incluyen la validación o invalidación de los primeros resultados y la introducción de instrumentos o métodos alternativos para obtener nueva información y conocimientos. Para que la investigación sea validada, las publicaciones científicas deben poder referirse a los datos disponibles y comparables desde una nueva óptica.

El concepto de datos que se puedan encontrar y sean accesibles, interoperables y reutilizables (principios FAIR) se está convirtiendo en el terreno común sobre el que se realizan y se realizarán las investigaciones. Entonces, ¿qué ocurre si los datos de investigación —cuya producción costó millones, si no miles de millones— dejan de estar técnicamente disponibles por falta de conservación?

La conservación de los datos digitales de investigación es esencial para la ciencia y el conocimiento. La definición de *planes de gestión de datos* se ha convertido en la norma para todo proyecto de investigación que solicite financiación, ya sea de su propia institución o de entidades externas. Sin embargo, no existe una definición de lo que hay que preservar, de cómo hay que preservarlo ni durante cuánto tiempo.

Desde el punto de vista archivístico, hay una creciente demanda de investigaciones relacionadas con los litigios y solicitudes de acceso a la información e investigación de la privacidad que pueden venir con la expectativa de que la información se libere digitalmente de manera oportuna. Con la plétora de versiones y copias creadas en un entorno digital, esto puede crear una importante carga administrativa para las instituciones de la memoria.

Preguntas esenciales para la selección

Las partes de los datos de la investigación que merecen ser conservadas a largo plazo como material del patrimonio deben definirse mediante un exhaustivo análisis crítico metodológico y profesional. Se trata de una metodología similar a la utilizada en los archivos. Se pueden plantear preguntas de este tipo:

- ¿Cuál es la calidad científica de los datos? ¿Cómo se crearon? ¿Por quién? ¿Cuál es su grado de apertura?

- ¿Existen restricciones de privacidad relacionadas con los datos y, en ese caso, durante cuánto tiempo son válidas esas restricciones?
- ¿Puede reproducirse fácilmente la experiencia que creó los datos, con un coste mínimo?
- ¿Los datos se crearon mediante una experiencia que se no puede reproducir o con un coste elevado?

Los Estados, las agencias de financiación y las organizaciones de investigación deben invertir en la preservación de los datos de investigación, ateniéndose a los principios FAIR.

La colaboración entre instituciones, a los niveles nacional o internacional, permite la construcción de infraestructuras sostenibles para gestionar el almacenamiento y la conservación de cantidades cada vez mayores de conjuntos de datos de investigación. También podrían considerarse modelos de custodia posterior en los que los datos permanezcan en una institución creadora. En algunos casos, la institución creadora puede conservar estos datos de forma más eficaz por sí misma, en lugar de enviarlos a una institución patrimonial.

Los bibliotecarios, los archiveros, los profesionales de la información, los técnicos de datos, los ingenieros, los informáticos y los investigadores tienen un papel que desempeñar, ampliando así el ámbito de las profesiones responsables de la preservación de nuestro patrimonio científico.

Apéndice 4: Análisis detallado de la evaluación y selección de los medios sociales

¿Qué son los medios sociales?

Los contenidos de los medios sociales son generados por los usuarios en un servicio web interactivo. Los contenidos pueden ser imágenes, textos, archivos de audio o lo que acepte la plataforma.

¿Cómo se identifican los registros de los medios sociales?

Cada registro de los medios sociales tiene un número de identificación o referencia para identificar un nuevo elemento en el sitio web del medio social. Los elementos de la estructura y el diseño que son iguales para todos los usuarios del sitio web social no son registros de medios sociales, pero deben conservarse con un sistema de archivo web como otros sitios web.

Hay muchas cuestiones importantes que se deben tener en cuenta a la hora de adquirir registros de medios sociales. Es decir, ¿es su institución la más indicada para reunir y conservar este material?

¿Cómo se trabaja con los proveedores de servicios? ¿Cómo se capturan los registros de los medios sociales? ¿Qué aspectos hay que tener en cuenta desde el punto de vista técnico?

Trabajamos con interfaces de programación de aplicaciones (API) y herramientas de descarga para obtener los ficheros con todos los registros. Hay muchas definiciones de lo que constituye una API; las más útiles son las propuestas por [Wikipedia](#) y la [Sociedad de Archiveros Americanos](#).

Los registros de las redes sociales se pueden capturar con API en la mayoría de los casos, pero en otros se necesitan técnicas más avanzadas como el *scraping*.

Estrategias y tipos de herramientas:

- **Scraping/crawling de la web:** Un *web scraper* (recopilación de información de forma automática) o *crawler* (rastreo de sitios web para organizar la información) es un *software* de navegación que descarga datos, como un navegador web que obtiene datos en bruto pero no formateados por defecto.
- **Pedidos API:** Obtención de datos con comandos de la consola del terminal que envían consultas a la interfaz de programación de aplicaciones (API) y reciben datos formateados con parámetros específicos.
- **Exportación de datos del servidor/perfil:** Formulación de peticiones formales a la web social y recepción de todos los datos de nuestros perfiles y los de los usuarios que gestionamos.

Los recursos que figuran a continuación ofrecen más información sobre este tipo de estrategias y herramientas que pueden emplearse para capturar medios sociales.

Consideraciones antes de capturar registros de los medios sociales:

- ¿Es su organización la más indicada para capturar esta información? En caso afirmativo, ¿hay asociados a los que deba hacer partícipes?
- ¿Cuánto quiere capturar? ¿Cuánto puede capturar? ¿Tiene suficiente espacio libre en sus servidores? ¿Tiene discos duros externos?
- ¿Tiene una estrategia de conservación para todos estos registros de los medios sociales?

Retos sociales de la recopilación de medios sociales

Si bien algunas plataformas de medios sociales son abiertas por defecto, otras se utilizan de forma más privada y están restringidas a un público determinado. Merece la pena analizar cómo puede afectar la recopilación de estos contenidos a las personas, las familias o las comunidades, y cómo el público al que van dirigidos sirve para determinar las decisiones de recopilación. Cabe plantearse las siguientes preguntas:

- ¿Consideran los usuarios los medios sociales como información pública o personal?
- ¿Se consideran los usuarios de las plataformas de medios sociales "editores" y esperan que sus contenidos sean susceptibles de ser recopilados por una agencia estatal?
- ¿Puede haber un marco de varios niveles en relación con la ética de la recopilación de información en los medios sociales? ¿Es necesario adoptar diferentes estrategias para diferentes plataformas?
- ¿Cuánto sabemos sobre los distintos usos que dan las personas a los medios sociales?
- ¿Cuándo es necesario pedir permiso para recopilar información y cuándo no? ¿Es posible solicitar un permiso retroactivo? ¿Qué tipo de acceso a estas colecciones puede proporcionarse en cada caso?

Retos legales de la recopilación a partir de medios sociales

- ¿Deben tratarse todos los medios sociales de la misma manera o hay diferentes estrategias según el uso que se haga de una plataforma de medios sociales en la sociedad?
- ¿Cómo es posible equilibrar el mandato de recopilación con la necesidad de proteger la privacidad? ¿Cómo se puede lograr garantizando al mismo tiempo que la institución es un repositorio de confianza para los datos personales?
- ¿Cuál es la posición de la jurisdicción local respecto al "derecho al olvido"?
- ¿Cómo se pueden gestionar los contenidos digitales de "jurisdicción mixta" (como el material de Twitter)?

Lecturas seleccionadas

- Corrado, E. M., Moulaison, H. L. (2017). *Digital Preservation for Libraries, Archives and Museums*. (2ª Ed.). Rowman & Littlefield.
- Russell, M. A., Klassen, M. (2018). *Mining the Social Web*. O'Reilly Media. (3ª Ed.)

Apéndice 5: Análisis detallado de la inteligencia artificial

Introducción

El uso de la inteligencia artificial (IA) se está convirtiendo en un pilar de las organizaciones de los sectores público y privado. Se considera un medio eficaz y eficiente para analizar grandes cantidades de datos estructurados (es decir, conjuntos de datos) y no estructurados (por ejemplo, documentos de procesamiento de textos, presentaciones o contenidos audiovisuales). La IA permite a las organizaciones obtener información acerca de sus datos que sería difícil de obtener mediante un proceso exclusivamente humano. Permite analizar grandes cantidades de datos para adoptar decisiones políticas, desarrollar tratamientos médicos, producir campañas publicitarias y mucho más. Los resultados de estas herramientas están facilitando el proceso de adopción de decisiones en varios sectores, lo cual genera registros que deben ser documentados y capturados.

Esta sección comenzará con una definición de la inteligencia artificial, a la que seguirá un análisis de sus componentes que deben ser evaluados y considerados durante el proceso de evaluación y selección. A continuación, se analizarán varios factores que afectan a la adquisición de registros de IA, como la repercusión de la asociación entre los sectores público y privado y la capacidad y los recursos de las instituciones del patrimonio documental para adquirir estos registros. La sección concluirá con un debate sobre la IA en los procesos de evaluación y selección.

Esta sección es un primer intento de definir un proceso en torno a la evaluación y selección de los registros de IA. Este proceso irá evolucionando a medida que los profesionales del patrimonio documental vayan adquiriendo nuevos conocimientos y experiencias.

Definición

La inteligencia artificial puede definirse de muchas maneras y no existe una única definición consensuada, pero a los efectos de este capítulo la definiremos como: "una rama de la informática que se ocupa de la simulación del comportamiento inteligente en los ordenadores; la capacidad de una máquina para imitar el comportamiento humano inteligente" (Marr, 2018). El concepto de IA puede dividirse en dos grandes grupos: supervisada y no supervisada.

Los algoritmos supervisados o de aprendizaje automático supervisado son aquellos en los que se entrena a un algoritmo para que identifique patrones utilizando datos de entrenamiento etiquetados; por ejemplo, nombres de personas o lugares. A continuación, se le proporcionan los datos brutos sin etiquetar para evaluar sus niveles de precisión ("precisión y exhaustividad", 2021).

Los algoritmos no supervisados o de aprendizaje automático no supervisado son aquellos en los que un algoritmo recibe datos no etiquetados y a continuación identifica patrones, utilizando probabilidades estadísticas incorporadas (es decir, inferencia bayesiana). Este

proceso, a diferencia del aprendizaje automático supervisado, se realiza con una mínima intervención humana ("aprendizaje no supervisado", 2021).

Hay muchas subcategorías y tipos diferentes de aprendizaje automático supervisado y no supervisado, como el aprendizaje automático profundo, la indexación semántica latente y el procesamiento del lenguaje natural. El objetivo de este capítulo no es definir y describir todos ellos, sino ofrecer al lector una visión general.

¿Cómo se evalúa y selecciona la IA?

Se deben tener en cuenta varios factores a la hora de valorar y seleccionar la inteligencia artificial. La IA, ya sea supervisada o no supervisada, va más allá de los resultados de los algoritmos. Hay muchos componentes diferentes en este tipo de documentos que pueden necesitar ser evaluados y capturados, como los ficheros, los datos (estructurados y no estructurados), la iteración del código y el código final junto con los resultados (es decir, las visualizaciones). Los profesionales también tendrán que evaluar el contexto de la creación, como el motivo del creador para crear y utilizar un algoritmo en la toma de decisiones, ya que esto afectará a lo que puede incluirse o quedar fuera del ámbito de la adquisición. Por último, también están las consideraciones legales que podrían afectar a lo que se puede adquirir. Esto se aborda en detalle en un capítulo anterior, pero en esta sección analizaremos las consecuencias de las asociaciones entre los sectores público y privado, en particular en lo que respecta al uso gubernamental de la IA en las decisiones políticas.

i) ¿Cuándo es la IA un registro?

Como se ha mencionado, pueden existir consideraciones legales que han de tenerse en cuenta en el momento de determinar si la IA constituye un registro que deba adquirirse, pero otras preguntas que los profesionales deberían plantearse para evaluar su valor a largo plazo son:

- ¿Repercutieron los resultados de la inteligencia artificial en la elaboración de políticas y en la aplicación de la política gubernamental? ¿Afecta esto a los ciudadanos? ¿Afecta a su capacidad para hacer valer y defender sus derechos?
- ¿Cambiaron los resultados de la inteligencia artificial la misión de una organización?
- ¿Influyó el uso de la inteligencia artificial en un proyecto clave de la organización? ¿Afectó a la forma en que se tomó la decisión?
- ¿El uso de la inteligencia artificial hizo que una organización o gobierno cambiara su metodología habitual de toma de decisiones?
- ¿Participaron múltiples entidades u organizaciones en el desarrollo y la formación de la IA? En caso afirmativo, ¿a quién pertenece el código de IA? ¿A quién pertenecen los datos de formación? ¿A quién pertenecen los resultados? ¿Son la misma entidad?

Básicamente, si la respuesta a los cuatro primeros puntos es "sí", entonces la IA constituye un registro de valor histórico perdurable que debe ser adquirido por una organización de patrimonio documental. La última pregunta trata de arrojar luz sobre los desafíos que surgen cuando hay más de una entidad implicada en la creación, el entrenamiento y el uso de la IA.

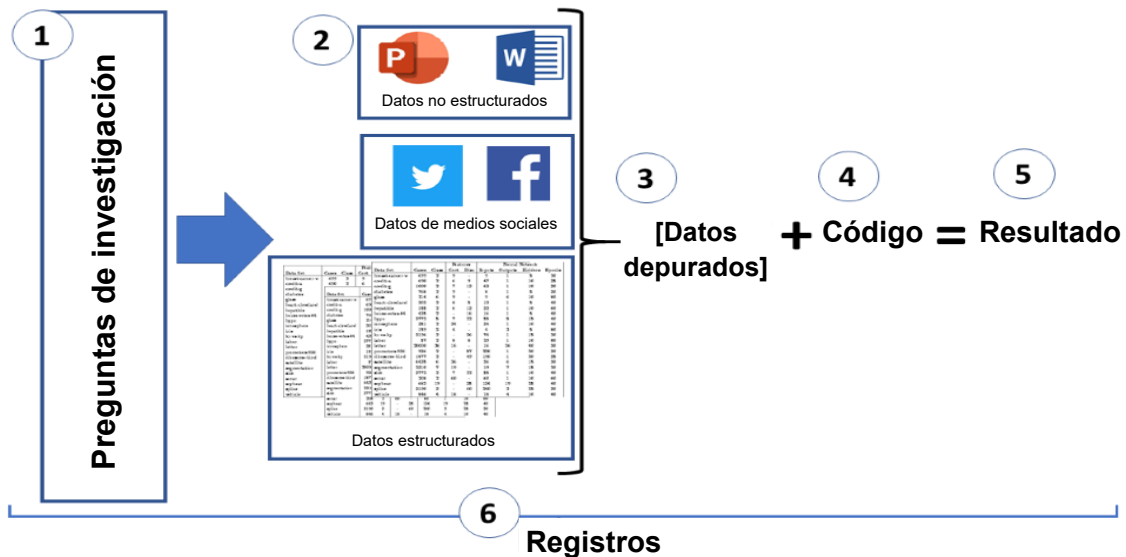
El último punto se analizará en profundidad en la sección *iii) Asociaciones entre los sectores público y privado*, pero podría repercutir de manera significativa en el proceso de selección.

Solo es posible evaluar el alcance de lo que puede adquirirse si se comprenden los diferentes componentes de la IA. Sin embargo, la amplitud y el volumen de los componentes identificados para la transferencia pueden superar la capacidad y las competencias de las instituciones de patrimonio documental. Las limitaciones a las que se enfrentan estas instituciones pueden requerir soluciones innovadoras y fuera de lo común para garantizar la captura y conservación de estos ficheros. Analizaremos algunas ideas en la sección *iv) Capacidad de las instituciones patrimoniales*.

ii) Componentes de la inteligencia artificial

Un registro de IA no es simplemente el resultado de un algoritmo, porque para entender cómo ha llegado a ese resultado, se necesitan los datos, los diarios y el código. En la figura 1.1 se describen los distintos componentes de un registro simple de IA, lo que significa que solo hay que capturar un fragmento de código junto con su documentación de apoyo (registros y datos). En esta figura se presupone que la organización que utiliza el algoritmo tiene el control de todo el proceso, desde la elaboración de las preguntas de investigación, pasando por la compilación y la depuración de datos, hasta el desarrollo y la iteración del código para llegar al resultado.

Figura 1.1 Componentes de un registro simple de inteligencia artificial



1) Preguntas de investigación

Cualquier iniciativa que utilice la inteligencia artificial tiene preguntas de investigación a las que debe dar respuesta. Estas preguntas serán la base para identificar los datos pertinentes y servirán de base para las decisiones relacionadas con la depuración de datos y las iteraciones en el desarrollo del código para llegar al resultado final. Un profesional deberá comprender las preguntas de investigación al principio del proceso de evaluación y selección, ya que pueden servir como base para determinar lo que será incluido o quedará fuera del ámbito de transferencia.

2) Datos

Se identificarán diferentes tipos de datos (estructurados, no estructurados y de medios sociales) que se reunirán para entrenar al algoritmo. Los datos estructurados son datos en forma de tabla (por ejemplo, una hoja de cálculo de Excel), mientras que los datos no estructurados son datos que no tienen un modelo predefinido (por ejemplo, un documento de tratamiento de textos o una presentación). Los medios sociales se pueden calificar como estructurados o no estructurados y, como tales, constituyen una categoría de datos independiente (Price, 2013). La selección de datos debe documentarse en un registro, y sirve como pista de auditoría para el proceso de desarrollo y entrenamiento de la IA. El registro debe contener información sobre los datos, como su origen, contenido, rango de fechas y conjunto de muestras.

3) Datos depurados

Será necesario agregar y depurar todos los datos procedentes de diferentes fuentes. El término *depuración* se refiere a los datos que no son relevantes para las preguntas de la investigación o que deben ser armonizados con otros datos (por ejemplo, cambiar de centímetros a metros) que necesitan ser introducidos en el algoritmo. La fase de depuración de datos es necesaria para facilitar la comparación y el análisis de la información por el algoritmo. Todas las decisiones adoptadas durante la fase de depuración de datos deben documentarse en un diario, lo que permitirá a los investigadores y a los informáticos reproducirlos para confirmar los hallazgos y las decisiones resultantes del análisis de la IA (Mackenzie, 2019).

4) Código

El código es una representación computacional y matemática de la respuesta a las preguntas de investigación. Permite al algoritmo interpretar y analizar los datos para llegar a una "respuesta". Los desarrolladores o investigadores desarrollan el código de forma iterativa para llegar a la mejor respuesta posible a las preguntas de investigación. Durante este proceso, debe evaluarse la exactitud (precisión y exhaustividad) para garantizar la integridad del resultado del algoritmo.

El proceso de desarrollo del código puede documentarse en los diarios y/o en un software concebido para auditar el proceso de desarrollo del código. Hay que insistir en la importancia de documentar este proceso para favorecer la reproducibilidad de los resultados del algoritmo, verificando la integridad de los hallazgos y, por extensión, las decisiones que se tomaron basándose en el resultado.

5) Resultado

El resultado es el fruto de los procesos computacionales aplicados por el algoritmo y puede presentarse de diferentes maneras, como visualizaciones o probabilidades estadísticas.

iii) Repercusión de las alianzas entre los sectores público y privado

En muchos casos, las iniciativas de IA, en particular en el sector público, consisten en asociaciones entre un organismo público y una empresa privada especializada en IA. Esto tiene mucha repercusión en el proceso de selección y está relacionado con el debate en torno a la propiedad intelectual y los derechos de autor que se ha planteado anteriormente en este documento. En esencia, los profesionales deben determinar:

- ¿A quién pertenece el código o algoritmo? Si es propiedad de la empresa, ¿cómo se documenta el resultado de la IA?
- ¿A quién pertenece el conjunto de datos del entrenamiento? ¿Y los datos posteriores introducidos en la IA?
- ¿A quién pertenece el resultado de la IA (por ejemplo, el conjunto de datos o la visualización)?

En teoría, los organismos públicos deben rendir cuentas por las decisiones que adoptan utilizando la IA, y aunque las consideraciones de propiedad intelectual deben tenerse en cuenta antes de su aplicación, puede no ser el caso cuando se trata de decisiones normativas. Esto plantea problemas éticos en cuanto a la rendición de cuentas, la transparencia y el futuro registro público, y pueden ser necesarias soluciones alternativas para documentar el uso de la IA en la toma de decisiones. Sería conveniente que los profesionales examinen los manifiestos de "responsabilidad algorítmica", como la Declaración de Montreal y la Declaración de la ACM.

iv) Capacidad de las instituciones patrimoniales

La cuestión de la capacidad (infraestructura, recursos y personal) se aborda en parte en el árbol de decisión para la evaluación y selección que figura en este capítulo, en particular en el apartado de "sostenibilidad", pero los profesionales deberían dejarse influir más por la importancia de la IA. La sostenibilidad será un reto constante, y la cuestión de cómo asimilar, preservar y elaborar estos materiales debe ser analizada en mayor profundidad por la comunidad del patrimonio documental. Por lo tanto, si bien será necesario evaluar la capacidad, la cuestión principal consistirá en decidir qué soluciones provisionales aplicar a corto plazo mientras la comunidad estudia otras estrategias más sostenibles a largo plazo.

Otras consideraciones

El uso de la IA en la evaluación y selección de los materiales de origen digital

La IA es un material que merece ser seleccionado y preservado, pero habida cuenta del volumen y la complejidad de los materiales digitales que deben ser evaluados para su selección en el futuro, los profesionales del patrimonio documental no pueden aplicar unilateralmente las técnicas tradicionales de valoración y selección. Tendrá que haber alguna forma de automatización en el futuro, y la profesión debe evaluar los puntos fuertes y débiles de estas tecnologías para entender en qué casos es mejor aplicarlas y cuándo. Hay instituciones del patrimonio documental que están experimentando con estas tecnologías para la evaluación y la selección.

Lecturas complementarias: La IA en la evaluación y la selección

- Caplan, R., Donovan, J., Hanson, L., y Matthews, J. (2018). Algorithmic Accountability: A Primer. *Data & Society*. https://datasociety.net/wp-content/uploads/2018/04/Data_Society_Algorithmic_Accountability_Primer_FINAL-4.pdf
- Chumtong, J., Kaldewey, D. (2017). Beyond the Google NGRAM Viewer: Bibliographic Databases and Journal Archives As Tools for Quantitative Analysis of Scientific and Meta-Scientific Concepts. *FIW Working Paper No 8*. <https://www.fiw.uni-bonn.de/publikationen/FIWWorkingPaper/fiw-working-paper-no.-8>
- Engin, Z., Treleaven, P. (2018, agosto). Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. *The British Computer Society*. <https://academic.oup.com/jnl/advance-article/doi/10.1093/comjnl/bxy082/5070384>
- Ertzscheid, O. (2017). *L'appétit des géants: pouvoir des algorithmes, ambitions des plateformes*. París: C&F.
- Comisionado de Información y Protección de la Intimidad. (2017). *Big Data, Artificial Intelligence, Machine Learning and Data Protection*. Londres: ICO. <https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf>
- LeSueur, A. (2016). Robot Government: Automated Decision-Making and its Implications for Parliament. [Borrador de capítulo para su publicación en *Parliament: Legislation and Accountability*]. Oxford, Reino Unido: Hart Publishing. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2668201
- Administración Nacional de Archivos y Registros. (2020, octubre). *Cognitive Technologies White Paper: Records Management Implications for the Internet of Things, Robotic Process Automation, Machine Learning and Artificial Intelligence*. Washington, D.C. <https://www.archives.gov/files/records-mgmt/policy/nara-cognitive-technologies-whitepaper.pdf>
- Archivos Nacionales del Reino Unido (2016) *The Application of Technology Assisted Review to Born-Digital Records Transfers, Inquiries and Beyond*. (2016). Londres. <http://www.nationalarchives.gov.uk/documents/technology-assisted-review-to-born-digital-records-transfer.pdf>
- O'Neill, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Nueva York: Crown Publishing.
- Padilla, T., Allen, L., Potvin, S., Roke Russey, E., Varner, S. (2017, marzo). Collections as Data. <https://doi.org/10.17605/OSF.IO/MX6UK>.
- Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., Stuart K. (2018, noviembre). More Human than Human? Artificial Intelligence in the Archive. *Archives and Manuscripts*, 47(2), 179-203.

- World Wide Web Foundation. (2017). *Algorithmic Accountability: Applying the Concept to Different Country Contexts*.
https://webfoundation.org/docs/2017/07/Algorithms_Report_WF.pdf

Apéndice 6: Gestión de los metadatos

Metadatos

Los metadatos suelen definirse como "datos sobre datos", lo cual, si bien es una definición acertada, no es muy precisa. En las instituciones que se ocupan del patrimonio, los metadatos necesarios deben considerarse como cualquier información (en formato digital o físico) que sea esencial para garantizar que el material digital que se conserva es, y siga siendo, accesible, inteligible y utilizable a lo largo del tiempo. Los metadatos proporcionan a la institución la información necesaria para acceder y preservar el patrimonio digital en el futuro.

Las instituciones que se ocupan del patrimonio suelen preservar tres tipologías clave de metadatos asociados al patrimonio digital que son cruciales para la conservación a largo plazo:

- Estructurales (necesarios para la capacidad técnica de lectura de contenidos digitales)
- Descriptivos (contienen información bibliográfica, archivística o museística contextual, que puede ser generada por el sistema o creada por profesionales del patrimonio, creadores de contenidos y/o usuarios)
- Administrativos (documentan la gestión de un objeto digital mientras forma parte de su colección)

Si el patrimonio digital es el "contenido", los metadatos proporcionan el "contexto".

Existen cinco requisitos funcionales básicos para los metadatos digitales:

Identificación: Los metadatos deben permitir que cada objeto digital sea identificado de forma única y sin ambigüedades. En general, esto requiere que se asigne a cada objeto un identificador único a nivel mundial.

Ubicación: Los metadatos deben permitir que cada objeto digital pueda ser localizado y recuperado. Se debe garantizar la validez a largo plazo de estos datos identificadores de ubicación, de tal manera que los objetos no se pierdan cuando se migran o se actualizan los sistemas.

Descripción: Se requiere una descripción del objeto digital para facilitar su recuperación e interpretación. Los metadatos descriptivos pueden clasificarse en dos categorías: datos sobre el contenido y datos sobre el contexto. Con frecuencia, los datos sobre el contenido de un objeto pueden ser recreados mediante exámenes y consultas. También son útiles como auxiliares para ayudar a descubrir recursos. Los datos sobre el contexto — dónde, cuándo y quién creó el objeto, para que se utilizó, o el lugar que ocupa en relación a un corpus general de materiales— son mucho más difíciles de recrear una vez que se han perdido.

Legibilidad: Los metadatos sobre la estructura, el formato y la codificación de los objetos digitales son indispensables para garantizar que sigan siendo legibles con el paso del tiempo. Este requisito funcional es especialmente importante para los objetos digitales, ya que estos no pueden ser leídos sin la mediación de una tecnología. Estos metadatos deben identificar las normas pertinentes y proporcionar referencias para la documentación técnica, los ficheros

de autoridad y otros materiales relacionados necesarios para una interpretación completa del recurso digital. Hay que prestar atención para garantizar que las múltiples capas de un objeto digital puedan ser interpretadas: desde el formato del fichero encapsulado hasta la representación y codificación de los datos en sí mismos.

Gestión de derechos: Los derechos, las condiciones de uso y las restricciones aplicables a cada objeto digital deben quedar registrados en los metadatos. Estos metadatos deben identificar las leyes y convenciones aplicables y aportar referencias sobre la documentación legal pertinente y los contratos, entre otras cosas, así como sobre los propietarios de los derechos.

Almacenamiento de los metadatos

Muchos formatos de ficheros digitales permiten la inserción de los metadatos dentro del fichero mismo. Esto tiene la ventaja de garantizar la permanencia del vínculo entre los datos y los metadatos. Sin embargo, los metadatos también deben ser almacenados independientemente del recurso digital que describen; esto es indispensable para cumplir los requisitos de funcionalidad mencionados más arriba. Un objeto digital codificado, por ejemplo, no puede ser leído si el código solo se encuentra insertado en el propio objeto.

Meta metadatos

Para poder establecer la fiabilidad y autenticidad de los metadatos se requiere información acerca de su origen y de cómo fueron compilados. La contextualización es indispensable para garantizar la futura recuperación y comprensión de la información digital.

Estos datos pueden incluir:

- ¿Cuándo fueron compilados los metadatos y quién lo hizo?
- ¿Los metadatos fueron recopilados automáticamente o manualmente?
- ¿Qué instrumentos y técnicas se utilizaron?

Apéndice 7: Referencias

Abelson, H., G. J. S., Sussman, J. (1985). *Structure and Interpretation of Computer Programs*. The MIT Press.

Cook, T. (1991). Many are Called but Few are Chosen. *Archivaria*, 32.
<https://archivaria.ca/index.php/archivaria/article/view/11759/12709>

Coalición de la Preservación Digital. (2015). *Digital Preservation Handbook* (2ª Ed.).
<https://www.dpconline.org/handbook>

Coalición de la Preservación Digital. (s.f.). *What is Digital Preservation?*
<https://www.dpconline.org/digipres/what-is-digipres>

Digital Strategy for the Library of Congress. (s.f.) La Biblioteca del Congreso. Consultado el 3 de diciembre de 2020 en <https://www.loc.gov/digital-strategy>

Harvey, R. (2006). Instalment on Appraisal and Selection. *DCC Digital Curation Manual* (versión 1). <https://www.era.lib.ed.ac.uk/bitstream/handle/1842/3331/Harvey%20appraisal-and-selection.pdf?sequence=1>

International Data Cooperation. (2020). *IDC's Global DataSphere Forecast Shows Continued Steady Growth in the Creation and Consumption of Data*.
<https://www.idc.com/getdoc.jsp?containerId=prUS46286020>

Koerbin, P., Webb, C., & Pearson, D. (2013). 'Oh, you wanted us to preserve that?!' Statements of Preservation Intent for the National Library of Australia's Digital Collections. *Magazine of Digital Library Research*, 19(1/2).
<http://www.dlib.org/dlib/january13/webb/01webb.html>

MacKenzie, R. J. (2019). Repeatability vs. Reproducibility. *Technology Networks*.
<https://www.technologynetworks.com/informatics/articles/repeatability-vs-reproducibility-317157>

Marr, B. (2018). The Key Definitions of Artificial Intelligence (AI) That Explain its Importance. *Forbes*. Consultado el 28 de enero de 2021 en <https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=dd7c4a64f5d8>

Biblioteca Nacional de Australia. (2003). *Directrices para la preservación del patrimonio digital*. UNESCO. Directrices para la preservación del patrimonio digital – UNESCO.
https://unesdoc.unesco.org/ark:/48223/pf0000130071_spa

Niu, J. (2012). An Overview of Web Archiving. *D-Lib Magazine*, 18(3/4).
<http://www.dlib.org/dlib/march12/niu/03niu1.print.html>

Niu, J. (2014). Appraisal and Selection for Digital Curation. *International Journal of Digital Curation*, 9(2), 65-82. <http://www.ijdc.net/index.php/ijdc/article/view/9.2.65>

Pearson, D. (2012, March 26-27). *The adventures of Digi: Ideas, requirements and reality* [presentación de la conferencia]. Future Perfect 2012, Wellington, New Zealand.
<https://www.nla.gov.au/content/the-adventures-of-digi-ideas-requirements-and-reality>

Precision and Recall. (2021). De la *Wikipedia*.
https://es.wikipedia.org/wiki/Precisi%C3%B3n_y_exhaustividad

Price, G. (2013). The Difference Between Structured and Unstructured Data in Social Media. *Social Media Today*. <https://www.socialmediatoday.com/content/difference-between-structured-and-unstructured-data-social-media>

Shustek, L. J. (2006). What Should We Collect to Preserve the History of Software? *IEEE Annals of the History of Computing*, 28(4), 110–112.

Slade, S., Pearson, D., & Knight, S. (2019). An introduction to digital preservation. In L. Elkin, & C.A. Norris (Eds.), *Preventive Conservation: Collection Storage* (págs. 809 a 829). Nueva York: Sociedad para la Preservación de la Historia Natural; Instituto Americano de Conservación de Obras Históricas y Artísticas; Institución Smithsonian; Programa de Estudios Museísticos de la Universidad George Washington.

Software Heritage. (2020). *SoftWare Heritage persistent IDentifiers (SWHIDs)* (versión 1.5).
<https://docs.softwareheritage.org/devel/swh-model/persistent-identifiers.html>

Tamaro, A. (2016). Heritage curation in the digital age: Professional challenges and opportunities. *International Information & Literary Review*, 48, 122-128.

UNESCO. (2003). *Carta sobre la preservación del patrimonio digital*.
http://portal.unesco.org/es/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html

UNESCO & INRIA. (2019). *Paris Call — Software Source Code as Heritage*.
<https://unesdoc.unesco.org/ark:/48223/pf0000366715>

UNESCO y Universidad de Columbia Británica. (2012, 26 a 28 de septiembre). *UNESCO/UBC Vancouver Declaration*. The Memory of the World in the Digital Age: Digitization and Preservation, Vancouver, British Columbia (Canadá).
http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_abc_vancouver_declaration_en.pdf

UNESCO, Universidad de Pisa e INRIA. (2019). *The Software Heritage Acquisition Process*.
<https://unesdoc.unesco.org/ark:/48223/pf0000371017/PDF/371017eng.pdf.multi>

Aprendizaje no supervisado. (2021), de *Wikipedia*.
https://es.wikipedia.org/wiki/Aprendizaje_no_supervisado

Werf, T. van der, & Werf, B. van der. (2014). *The paradox of Selection in the Digital Age* [presentación]. Sesión pública de la UNESCO, IFLA WLIC 2014, Lyon (Francia).
<http://library.ifla.org/id/eprint/1042>