"Los Archivos Electrónicos y el Proyecto InterPARES"[1],[2]

Conferencia
Dirección General del Archivo Nacional
San José de Costa Rica
Lunes 23 de agosto de 2010

La súbita incorporación de las tecnologías de la información y las comunicaciones (TIC) en prácticamente todas las actividades que el ser humano realiza tanto en su entorno laboral como social tiene beneficios indiscutibles en tanto que se rompen las barreras físicas para la adquisición de conocimiento. No obstante, prevalecen serias desigualdades y baste un ejemplo de ello en México, según cifras del Instituto Nacional de Geografía e Informática del 25.7% de hogares con computadora, solo el 13.5% tiene conexión a internet, de éstos el 53% de los hogares manifiesta que no cuenta con el mismo por falta de recursos económicos.[3] Cierto es que existe una porción de la población versados en el uso y manejo de información mediante el uso de las TIC, pero con seguridad, también se encuentra un importante número de personas analfabetas o analfabetas funcionales respecto de las TIC. Por otra parte, la información en la web aún requiere de estrategias de mejora para su fácil acceso, los motores buscadores por lo general carecen de la estructura organizacional de la información y en la mayoría de los casos, ante una consulta, arrojan un cúmulo de información sin orden alguno.

---

[1] Alicia Barnard. Es actualmente consultora en Archivos y participante del Proyecto InterPARES. barnard.alicia@gmail.com

[2] Esta presentación, con adaptaciones y ajustes, tiene como base la conferencia presentada durante los festejos del X Aniversario del Archivo Histórico de la Universidad de Sonora "X Años pro la Memoria Universitaria 1999-2009", Hermosillo, Sonora México, marzo 25, 2009. También publicado en: Boletin: Expedientes Universitarios. Nueva Época. 7: pp 136-145, junio, 2009.

[3] Ver estadísticas de INEGI disponibles en http://www.inegi.gob.mx/inegi/default.aspx?c=8823&s=est

La irrupción de las TIC en oficinas de gobierno, instituciones académicas, empresas también ha traído como consecuencia desequilibrios, mientras los expertos en las TIC han desarrollado un cúmulo de herramientas que facilitan el desarrollo de las actividades, la generación de información, el acceso a ésta que por derecho tienen los ciudadanos, la gestión de trámites, la transparencia y rendición de cuentas de las acciones, acciones destinadas al manejo y control de ese cúmulo de información que se han relegado. Este tipo de información, al derivarse de una función o atribución, puede tratarse de un archivo digital[4] que debería estar sujeto a las prácticas archivísticas para su control a partir de su creación, y durante su uso, mantenimiento y preservación en el largo plazo cuando los mismos tienen valores históricos o son de importancia vital para las organizaciones. Así, mientras estos expertos en tecnologías de la información desconocen acerca de la práctica archivística que implica entender acerca de las agregaciones de documentos de archivo, su contexto y sus valores para ser manejadas de forma óptima al paso del tiempo a los archivistas esta irrupción de las TIC nos encontraron sin el conocimiento de útil para entender el comportamiento de los documentos de archivo digitales y contender con los riesgos que su fragilidad y volatilidad representan.

Se podría decir que el estudio sobre los documentos de archivo que nacen digitalmente tiene poco más de dos décadas. El saber si en los ambientes electrónicos existían documentos de archivo quedó claro cuando las Naciones Unidas en 1990 a través del Comité para la Coordinación de Sistemas de Información (ACCIS por sus siglas en

[4] Nota. En el contexto de este trabajo los conceptos "documento de archivo", "documento archivo electrónico" o "documento de archivo digital", o en su forma genérica "archivos digitales" "archivos electrónicos serán utilizados indistintamente. En el diccionario de la base de datos de terminología de InterPARES se define al documento digital de archivo "*como aquel que ha sido capturado y fijado para su almacenamiento y manipulación en un sistema de computadora y que requiere el uso del sistema para ser inteligible por una persona".* Otra definición del mismo diccionario dice que se trata *de "un documento de archivo cuyo contenido y forma están codificados mediante el uso de valores numéricos discretos (tales como valores binarios 0 y 1), en lugar de un espectro continuo de valores (tales como aquellos generados por un sistema análogo)".* También en la misma base se localiza la definición de documento de archivo electrónico: *Documento de archivo analógico o digital que es transportado por un conductor eléctrico y requiere del uso de equipo electrónico para ser inteligible por una persona*. InterPARES 2 Terminology Database. http://www.interpares.org/ip2/ip2_terminology_db.cfm (consulta, julio, 2010)

inglés), realizó uno de los estudios más completos en la materia. Para ello consideró las prácticas de la administración de documentos y la archivística y como resultado en el reporte de este estudio, llevado a cabo en las 18 organizaciones de las Naciones Unidas, se estableció que había que "distinguir entre el documento de archivo y otro material".[5] Las recomendaciones de este estudio fueron revisadas por investigadores de la Universidad de Columbia Británica en Vancouver, Canadá, quienes iniciaron un proyecto de investigación (conocido como Proyecto UBC) cuyo objetivo primordial, en principio, fue el de establecer lo que era un documento de archivo en principio y cómo podía reconocerse en el ambiente electrónico. Para ello, la investigación estuvo orientada bajo el supuesto de que los métodos y controles para archivos electrónicos solo podían provenir de principios y conceptos ya utilizados en archivos en ambientes tradicionales, para lo cual se utilizaron los conceptos de la diplomática como ciencia dedicada al estudio del documento de archivo en su forma individual y a la ciencia archivística dedicada al estudio del documento de archivo como agregaciones. El grupo de investigación en conjunto con otros expertos del Departamento de Defensa de los Estados Unidos de Norteamérica y de los Archivos Nacionales de Estados Unidos tuvieron como resultado el primer estándar de criterios para los archivos electrónicos la DOD.5015.2. [6]

El Proyecto InterPARES

Es entonces el Proyecto UBC el precursor del Proyecto Internacional Investigación sobre Documentos de Archivo Permanentes en Sistemas Electrónicos (InterPARES, por sus siglas en inglés, acrónimo que también significa en inglés "entre pares") con sede en la ya mencionada Universidad la Columbia Británica. InterPARES tiene su inicio en 1999 y, por

[5] Luciana Duranti, Terry Eastwood and Heather MacNeil (2002) : Preservation of the Integrity of Eletronic Records. Kluwer Academic Publishers. Dodrecth, pp. 1.
[6] Ibid 4, pp. 4-5

ahora continuará hasta 2012, el mismo se ha convertido en el proyecto de investigación para la preservación de archivos digitales más antiguo, si es que se puede llamar antiguo a un periodo de 11 años. El Proyecto se ha orientado al desarrollo de conocimiento para preservar, documentos auténticos, creados o mantenidos en forma digital, así como a proporcionar bases para estándares, políticas, estrategias, planes de acción que permitan asegurar la duración de este tipo de materiales y la habilidad de sus usuarios para confiar en su autenticidad.

El Proyecto InterPARES se ha desarrollado en tres fases. Así, InterPARES 1 inició en 1999 y concluyó en el 2001. Durante esta etapa se llevaron a cabo estudios dedicados a desarrollar el conocimiento teórico metodológico para la asegurar la autenticidad de archivos creados o mantenidos en bases de datos y sistemas electrónicos de gestión documental que ya se encontraban en archivos históricos, por lo que se tuvo como punto de vista la visión del preservador. A través de las investigaciones de InterPARES 1, se llegó a las siguientes conclusiones:

a) Para contar con archivos fiables y precisos y mantener su autenticidad es necesario establecer requisitos desde la creación de los documentos;

b) No es posible preservar documentos de archivo digitales sino la habilidad para reproducirlos, y

c) La necesidad de que el preservador se involucre con documentos de archivo al principio de su ciclo de vida a fin de validar que aquellos seleccionados para su preservación permanente sean copias autenticas del creador.

InterPARES 2 inició en el 2002 y concluyó en el 2006. Además de continuar con temas de autenticidad se realizaron investigaciones sobre fiabilidad y precisión durante el ciclo de vida de los documentos de archivo, desde su creación hasta su preservación permanente. Los casos de estudio estuvieron orientados a documentos de archivo que se producen en ambientes digitales complejos durante el desarrollo de actividades artísticas, científicas y de e-gobierno. Durante esta fase, gracias a las gestiones de la Dr. Luciana Duranti, Directora del Proyecto InterPARES, es que la UNESCO otorga financiamiento para que archivistas de Latinoamérica recibieran capacitación sobre archivos electrónicos y su preservación en el largo plazo, las profesionales seleccionadas procedían de Argentina, Brasil, Cuba, México y Perú.

El proyecto InterPARES se sustenta en principios, mismos que han resultado de las propias investigaciones realizadas durante IP1 e IP2.

- La tecnología no puede determinar la solución para la creación de documentos de archivo digitales fiables y precisos o para su preservación auténtica en el largo plazo: **los requerimientos organizacionales definen el problema y los principios archivísticos deben establecer el buen uso y adecuación de cada solución técnica.**

- Las soluciones para los retos en los documentos de archivo digitales son intrínsecamente **dinámicas y específicas a las situaciones culturales, disciplinarias, administrativas y legales.**

- La preservación es un **proceso continuo que inicia con la creación de los documentos de archivo.**

- Se debe contar la capacidad para **asumir la fiabilidad de los documentos de archivo hasta que lo contrario sea establecido.**

- Se debe contar con la capacidad de **inferir autenticidad sobre la base de las circunstancias de la creación, mantenimiento y preservación de los documentos de archivo.**[7]

La fase 3 de InterPARES inicia en 2007 y concluirá en 2012. En esta fase el proyecto se construye sobre hallazgos de las primeras etapas, así como de otros proyectos de preservación a nivel internacional. Los componentes de proyecto son la investigación, mediante la puesta en práctica de la teoría y el conocimiento adquirido en las fases previas, esto mediante casos de estudio a realizar en unidades de archivo de tamaño mediano o pequeño; la educación y capacitación a través del desarrollo de módulos para capacitación, educación continua y curricula académica; así como actividades de divulgación de los resultados del proyecto. En su organización se contemplan los TEAM's (en español Elaboración Teórica sobre Gestión Documental, el acrónimo en inglés también significa equipo) de los cuales hasta ahora se han incorporado 16.

México se ha integrado a esta iniciativa de investigación internacional y la dirección del TEAM México está a cargo del Dr. Juan Voutsáss, Académico de la Universidad Nacional Autónoma de México (UNAM) e Investigador del Centro Universitario de Investigaciones Bibliotecológicas (CUIB) de la citada Institución. Hasta ahora participan en el TEAM, funcionarios e investigadores del Banco de México, del Tribunal Electoral del Poder

[7] Barnard,Alicia. Archivos Electrónicos (2008) . Hallazgos del Proyecto InterPARES. 2nda. Jornada del Patrimonio Documental Universitario. Dirección del Archivo General, Universidad de Guanajuato, Gto. Agosto, 29.

Judicial de la Federación y de la propia UNAM. De las actividades a desarrollar a través del TEAM México, se espera: contar con lineamientos o mecanismos específicos para la preservación de los archivos digitales de casos de estudio en particular, llevar a cabo actividades intensivas de divulgación del proyecto, así como prepara material documental útil para la capacitación y formación profesional. De lo anterior, hasta ahora se han llevado a cabo trabajos intensos en cuanto a la traducción de materiales y actividades de divulgación, se ha concluido la elaboración de políticas de preservación del Banco de México, mismas que se encuentran en periodo de comentarios por el resto de los miembros de InterPARES 3; asimismo se han impartido dos cursos sobre archivos electrónicos y su preservación y se espera iniciar un diplomado proximamente.[8]

Para conocer brevemente acerca de algunos hallazgos del proyecto InterPARES, a continuación se comenta sobre la naturaleza de los documentos de archivo digitales mismos que forman parte de la base de datos terminológica desarrollada durante InterPARES 2,[9] y brevemente se mencionan algunos de los productos desarrollados en InterPARES 1 y 2, los cuales sirven de base para los estudios de IP3 y que pueden ser útiles para quienes tienen a su cargo o tendrán en el futuro cercano archivos digitales.

Conceptos y definiciones acerca del documento de archivo digital

La mencionada base de datos está compuesta por un glosario, un diccionario y ontologías. El glosario viene a ser una lista de términos y definiciones que son base para los estudios de InterPARES principalmente se comentan los conceptos y definiciones

[8] Para mayor información sobre los trabajos del l TEAM México ver http://www.interpares.org/ip3/ip3_index.cfm (consulta, julio 2010)

[9] InterPARES 2 Terminology Database. http://www.interpares.org/ip2/ip2_terminology_db.cfm, (consulta, julio 2010) . Los avances de traducción al español se localizan en http://www.interpares.org/ip3/ip3_documents.cfm (consulta, julio 2010)

sobre el documento de archivo digital, así como aquellos relacionados con la confianza en los mismos.

Así, se define a un **documento de archivo** como aquel creado (elaborado o recibido) y salvado para una acción futura o como referencia por una persona física u organización en el curso de una actividad práctica como un instrumento y derivado de esa actividad.

Uno de los factores relevantes dentro de los archivos en formato digital es la confianza que se puede tener de los mismos ello implica que los mismos sean fiables, auténticos y exactos. Así en un documento de archivo su **fiabilidad** como declaración de un hecho se da cuando el mismo puede sostener al hecho del que es relativo y es establecida examinando la **completitud** en la forma del documento de archivo y el nivel de control ejercido durante su **proceso de creación** La **autenticidad** es define, entonces, como la cualidad que tiene un documento de archivo para ser lo que pretende ser y estar libre de alteración o corrupción. Así, para presumir que un documento de archivo digital es auténtico al paso del tiempo se requiere **proteger su integridad**, esto es, que debe estar completo y sin alteraciones, así como **mantener su identidad** para lo cual es necesario preservar las características que lo identifican de forma inequívoca y lo distinguen de cualquier otro documento de archivoy sobre éste define a la **precisión** se entiende al grado que alcanzan los datos, la información, los documentos o los documentos de archivo en cuanto a su exactitud, su forma correcta, su veracidad, el estar exentos de error o distorsión y el ser pertinentes a la materia de que tratan.[10]

[10] Ibid. 4. Para mayor información sobre la confianza en un documento de archivo revísese la ontología correspondiente que se localiza en la base de datos de terminología de InterPARES.

Las investigaciones de InterPARES, han tenido como resultado el refinamiento de conceptos, enseguida se hace referencia a las características y tipos de archivos digitales y de aquellos sistemas que contienen información susceptible de considerarse como un documento de archivo digital en potencia. Así, un documento de archivo debe contar con:

Contenido estable y forma fija. Se entiende por contenido estable cuando éste no cambia o cuando los cambios están determinados conforme a reglas establecidas, por contar con una **variabilidad vinculada**, es decir los cambios a la forma y/o contendido de un documento digital de archivo están limitados y contralados por reglas preestablecidas, a fin de que la misma pregunta, solicitud o interacción siempre genere el mismo resultado. **La forma fija** es la cualidad del documento de archivo que asegura que su contenido permanece completo y sin alteraciones, éste puede ser presentado en la pantalla en diferentes formas en una serie limitada de posibilidades, entonces se tiene una presentación documental diferente del mismo documento de archivo almacenado con **contenido estable** y **forma fija** (ejemplo, datos estadísticos vistos como un pastel o como barras o tablas).

En el documento de archivo digital se deben identificar los siguientes contextos: el **contexto jurídico-administrativo** corresponde al sistema legal y organizacional a los cuales pertenece el área creadora de los documentos de archivo. El **contexto documental** se refiere a los fondos a los cuales pertenece el documento de archivo y su estructura interna. Por lo que hace al **contexto procedimental**, éste se trata del procedimiento dentro del cual el documento de archivo se crea. En cuanto al **contexto de procedencia,** el mismo corresponde a la misión o mandato y funciones de la organización del creador. Por último, el **contexto tecnológico** se refiere al entorno o plataforma tecnológica dentro del cual se elaboró el documento de archivo.

También el documento de archivo digital debe tener menos las siguientes personas involucradas en su creación:

- Un autor que es la persona con autoridad para emitir un documento
- Un destinatario a quien se envía el documento
- Un escritor responsable de elaborar el documento (si es diferente al autor)
- Un creador identificado la persona (física o moral) que produce los fondos archivísticos al cual pertenece el documento de archivo (si es diferente al autor o al escritor).
- Un originador o persona a la que pertenece la dirección electrónica o el espacio en el cual el documento de archivo fue generado (desde donde es transmitido o en el cual se compila y guarda).

El documento de archivo digital deberá además participar o sustentar **acciones** derivadas de un procedimiento o como parte de la toma de decisiones. En cuanto a los elementos de **forma,** se trata de la parte constitutiva de la forma documental del documento de archivo, que es visible en el mismo. Los elementos documentales pueden ser extrínsecos como un sello o intrínsecos como la identificación del autor o del creador. Los **metadatos, son los atributos** de los documentos de archivo que demuestran su identidad e integridad tales como los nombres de quienes intervienen; acción o materia, fecha de creación y transmisión, clasificación, nombre de la institución, indicación de modificaciones técnicas o de anotaciones. Por su parte, los **componentes digitales** son entidades u objetos que pertenecen a uno o más documentos digitales de archivo, incluyendo los metadatos necesarios para ordenar, estructurar o hacer manifiesto el contenido, que requieren una medida específica de preservación, por ejemplo un documento de archivo puede tener texto en Word (cuya fuente o tipo de letra es un

componente digital), cuadros en Excell y además se cuenta con firma electrónica, cada uno es un componente digital.

Durante las investigaciones de InterPARES 2, se encontró que existen documentos de archivo **almacenados**, es decir, los componentes digitales utilizados para reproducir uno o más documentos de archivo, que incluye los datos para ser procesados a fin de que el documento de archivo se manifieste (datos de contenido y datos de forma) así como las reglas para procesar estos datos, incluyendo aquellas variaciones para su habilitación (datos de composición). Por otra parte, el documento de archivo **manifestado** es la visualización o materialización de éstos de forma apropiada para ser vistos por una persona o sistema. En ocasiones un documento de archivo no tiene un correspondiente almacenado, el mismo es tan solo una forma recreada de los datos de contenido fijados cuando la acción del usuario los asocia específicamente a datos de forma y datos de composición (ejemplo un documento de archivo producido desde una base de datos relacional).

También se halló que existen dos tipos principales de documentos digitales de archivo. **Estáticos.** Aquellos que no tienen posibilidades para cambiar su contenido manifiesto o forma más allá de abrir, cerrar y navegar, como el correo electrónico, reportes, video o música. **Interactivos**. Estos archivos presentan un contenido o una forma variable o ambos; las reglas que regulan su contenido y forma de presentación pueden ser fijas o variables. En un archivo digital interactivo la **forma fija** es almacenada de tal forma que el mensaje que contiene puede ser presentado con la misma presentación documental que tenía en la pantalla cuando fue salvado por primera vez (una presentación diferente sería de Word a PDF).

Por otra parte, en los sistemas se han identificado entidades digitales que por sus características contienen documentos de archivo potenciales.

- Entidades donde la variación se debe a los datos que cambian frecuentemente (el diseño permite actualización, cambio o alteraciones; también es posible la recolección de datos por parte de los usuarios o las interacciones o acciones de éste, o usos de estos datos para determinar subsecuentes presentaciones).

- Entidades donde la variación se debe a datos recibidos de fuentes externas y no son almacenadas dentro del sistema.

- Entidades producidas en el marco de aplicaciones dinámicas de cómputo que seleccionan diferentes grupos de reglas para producir documentos, dependiendo de la entrada (input) del usuario, fuentes de datos de contenido y características del contenido (sitios del estado del tiempo).

- Entidades producidas por computación evolucionaría donde el software que los genera pueden cambiar de forma autónoma (programación y modelaje de mercados financiero; sitios juegos educativos).[11]

Durante InterPARES 1 y 2, además de un número importante de casos de estudio se han publicado dos libros con los resultados de las investigaciones, se desarrolló una base de datos terminológica que incluye glosario, diccionario y tres ontologías, el Modelo de la

[11] Duranti, Luciana. Continuity and transformation in the role of the archivist: The Findings of the InterPARES Project. Memorias del Seminario Internacional de Archivos y Documentos Electrónicos, Universidad de La Salle Bogotá, marzo 12, 2009 (CD-ROM)

Cadena de Preservación es una propuesta para la integración de los sistemas de creación, mantenimiento y preservación.

De igual manera, dentro de los productos de InterPARES 2 se tienen dos de gran utilidad, mismos que ya han sido traducidos al español por el doctor Voutssás, director del TEAM México estos son La *Guía para la Creación y Conservación de Materiales Digitales. Lineamientos para Individuos*[12] y la *Guía para la Preservación de Archivos Digitales. Lineamientos para las Organizaciones*[13].

En la *Guía para la Creación y Conservación de Materiales Digitales* se reúne un conjunto de recomendaciones útiles que incluyen:

- Selección de hardware, software, formatos de archivo, que ofrecen una mejor oportunidad de asegurar que los materiales digitales permanecerán fácilmente accesibles en el tiempo.
- Aseguramiento en cuanto a la estabilidad y fijación de los materiales digitales mantenidos como documentos de archivo, tanto en su contenido como en su forma. Identificación adecuada de los materiales digitales.
- Garantía de que los materiales digitales cuenten con información que ayudará a verificar su integridad.
- Organización de los materiales digitales en agrupaciones lógicas
- Uso de técnicas de autenticación que promuevan el mantenimiento y preservación de los archivos digitales

[12] Disponible en: http://www.interpares.org/ip2/guia_del_creador/guia_del_creador.html (consulta , julio 2010)
[13] Disponible en: http://www.interpares.org/ip2/guia_del_preservador/guia_del_preservador.html (consulta, julio 2010)

- Protección de los materiales digitales para evitar acciones no autorizadas
- Protección de los materiales digitales para evitar de pérdida accidental o corrupción.
- Pasos en contra de la obsolescencia de hardware y software
- Temas acerca de la preservación en el largo plazo

Por su parte, en la *Guía para la Preservación de Archivos Digitales. Lineamientos para las Organizaciones* se incluyen los siguientes temas:

- **Manejar la cadena de preservación**. involucra aspectos relacionados con requisitos, diseño e instrumentación del esquema de la cadena de preservación que comprende los elementos de la política, estrategia, metodología, entre otros, para manejar los documentos digitales de archivo, tales como el establecimiento del alcance y objetivos en los programas de preservación digital, adquisición de recursos, enfoque sobre archivos digitales, desarrollo de procedimientos sobre la transferencia, mantenimiento y reproducción de los documentos de archivo, descripción archivística e instrumentación de estrategias de mantenimiento.
- **Valoración de los documentos para su preservación permanente**. Incluye llevar a cabo la valoración al principio, es decir, a partir de la creación de documentos; ubicar de dueños múltiples; evaluar la autenticidad conforme a los requisitos arriba mencionados, así como documentar la misma; monitorear los documentos identificados para su preservación en el largo plazo a fin de revisar posibles cambios en cuanto a actualización de software o hardware, cambios dentro de las instituciones, así como la actualización de la valoración ante posibles cambios o disposiciones legales; identificación de los componentes digitales y determinación de la viabilidad de preservación.

- **Transferencia de archivos digitales**. Implica el desarrollo de un plan compartido de transferencia, cumplir con procedimientos estandarizados, conservar el formato lógico más antiguo disponible, evitar duplicados y documentar todo el proceso de transferencia.
- **Preservación de los archivos transferidos**. Descripción de documentos de archivo en cuanto a sus contextos, identificar ramificaciones legales derivadas de las acciones de preservación, confirmar la efectividad de la estrategia de preservación, mantener un almacenamiento adecuado.
- **Dar acceso a los documentos de archivo**. Comprende explicar cómo se elaboran las copias, explicar cuáles son los requisitos de acceso.

De lo anterior se puede concluir:

La urgente necesidad de que las instituciones dentro de sus programas y presupuesto incorporen a la preservación de los archivos digitales como una activo de la misma ya que éstos son la materia prima que les da sustento y razón de ser.

La relevancia que debe darse a la investigación sobre la preservación de materiales digitales como mecanismo para instrumentar planes y programas específicos de acuerdo con las características propias de los sistemas donde estos materiales se crean y mantienen.

La imperante de que los profesionales de las TIC y otros profesionales de diferentes campos conozcan más acerca de la archivística y de que los archivistas aprendamos acerca de las TIC, ya que el diálogo fluido es uno de los mecanismos importantes en la preservación de archivos digitales.