

In dubio pro archivo:

Desarrollo de un proceso de ingesta y pista de auditoría, conforme con las regulaciones españolas acerca de e-administración, y haciendo uso de metodologías ágiles

Alejandro Delgado Gómez

Archivo Municipal de Cartagena

1. Introducción

En el presente texto exploramos el caso de una implantación específica del proceso de ingesta, junto con su pista de auditoría, basado en el modelo de referencia OAIS (Open Archival Information System) en el Archivo Municipal del Ayuntamiento de Cartagena. En primer lugar, contextualizamos la necesidad de este proceso de ingesta, codificando los requisitos legales en el Derecho español. En segundo lugar, explicamos los motivos para priorizar el proceso de ingesta. Seguidamente, describimos las tecnologías utilizadas por los miembros del proyecto, así como los motivos para elegir éstas y no otras. Por último, explicamos con más detalle los resultados de este proyecto, vinculado a, pero independiente de, el proyecto de desarrollo de un OAIS.

2. Requisitos legales

Las regulaciones españolas no son muy extensas con respecto a los archivos digitales. Sin embargo, codifican en algunos puntos la naturaleza sustancial de los mismos para las administraciones públicas. Describimos estos puntos en lo que sigue.

La Ley 11/2007, de 22 de junio, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos, meramente establece como obligatoria la necesidad de conservar los documentos electrónicos como tales documentos electrónicos, aunque no la necesidad de hacerlo en un archivo (Capítulo 4, art. 31.2).

Sin embargo, el Real Decreto 4/2010, de 8 de enero, que regula el Esquema Nacional de Interoperabilidad en el Ámbito de la Administración Electrónica, establece, en su artículo 21.1.j, como condición para el mantenimiento en el tiempo de los expedientes y documentos, la transferencia, cuando sea necesario, de éstos entre diferentes depósitos electrónicos con la finalidad de conservarlos y para poder asegurar su recuperación a medio

y largo plazo. Es decir, el ENI parece sugerir que, como en el mundo físico, los expedientes electrónicos tienen que ser transferidos a una autoridad archivística.

El mismo texto establece, en su artículo 21.2, la naturaleza obligatoria de un depósito digital, que crearán las administraciones públicas, con funcionalidad complementaria y equivalente a los archivos convencionales, y destinado a cubrir todo el ciclo de vida del documento.

Algunas de las Normas Técnicas de Interoperabilidad que desarrollan el Esquema Nacional de Interoperabilidad, de manera específica la Norma Técnica de Interoperabilidad de Política de Gestión de Documentos Electrónicos y la Norma Técnica de Interoperabilidad de Política de Firma y de Certificados Electrónicos de la Administración, también establecen la necesidad de conservar los documentos digitales a lo largo del tiempo y en archivos digitales, aunque no lo hacen de forma detallada. A pesar de llamarse normas técnicas, fueron promulgadas como resoluciones, una fórmula legal que las convierte en obligatorias.

Mientras escribimos este texto, dos proyectos de ley, la Ley de Procedimiento Administrativo y la Ley de Régimen Jurídico del Sector Público, también prevén archivos digitales obligatorios para cada una de las administraciones públicas en España.

El Poder Judicial y algunas Comunidades Autónomas, por su parte, han promulgado sus propias regulaciones, aunque quedan fuera del alcance de este texto.

A nivel local, el Ayuntamiento de Cartagena ha publicado dos herramientas reguladoras, la Ordenanza sobre Administración Electrónica de 8 de marzo de 2010 y la Política de Gestión de Documentos Electrónicos de 26 de julio de 2014. El Reglamento del Archivo Municipal se encuentra en fase de revisión para su adaptación a la nueva realidad.

Ninguno de estos puntos fuerza a las administraciones públicas a adoptar una aproximación específica. Es decir, no es obligatorio un OAIS.

3. El problema

Sin embargo, el Archivo Municipal tomó la decisión de desarrollar un proceso de ingesta basado en OAIS sobre la base de las regulaciones descritas anteriormente, para asegurar un cumplimiento preciso con el Derecho. Además, como hemos explicado en otros lugares (Delgado Gómez 2007, 2010, 2009, 2013), aunque la situación ideal es aquella en la que el sistema de conservación se encuentra integrado en el sistema de producción, el comportamiento de las burocracias aconseja una cierta separación, puesto que un sistema de conservación fuertemente vinculado a un sistema de producción podría poner en riesgo las propiedades de los documentos: autenticidad, integridad, fiabilidad y usabilidad. Por tanto, como premisa, se tomó la decisión de mantener separados los sistemas de

producción y de conservación, aunque el segundo insertaría, de manera no intrusiva, los controles archivísticos y de gestión documental en el primero.

El motivo para adoptar esta decisión encontró apoyo en un par de artículos, todavía no mencionados, acerca de los requisitos que una transferencia tiene que satisfacer para ser aceptada por el receptor, en este caso el Archivo. La Norma Técnica de Interoperabilidad de Documento Electrónico establece, en su artículo VII.5, que en el caso de intercambio de documentos electrónicos entre diferentes administraciones públicas que impliquen la transferencia permanente de responsabilidad sobre la gestión de los documentos, la organización que transfiere supervisará la autenticidad y la integridad de los documentos en el momento en el que el intercambio tenga lugar. La Norma Técnica de Interoperabilidad de Expediente Electrónico incluye un artículo, el V.6, en el mismo sentido. Bien es cierto que en el caso del Ayuntamiento de Cartagena la transferencia al Archivo no implica un cambio de responsabilidad entre administraciones, sino entre órganos de la misma administración; y que el riesgo de sospecha queda minimizado por el hecho de que el Archivo ha sido partícipe muy activo en el diseño de la férrea aplicación de gestión de expedientes en uso; pero, puesto que la posibilidad de error, intencionado o no, existe, la utilización de los citados artículos para establecer una capa adicional de garantía no nos parece descabellada.

Tengamos en cuenta que, desde el punto de vista del archivero, un grave problema reside en el hecho de que las regulaciones acerca de e-administración no incluyen un procedimiento sancionador; es decir, los archivos tienen que confiar en que el sistema de producción siempre transferirá documentos generados de manera adecuada, confianza que a lo largo de la historia ha sido quebrantada en demasiadas ocasiones. Para evitar este margen subjetivo de confianza, el Archivo Municipal desarrolló dos mecanismos, uno de ellos procedimental y el otro tecnológico. El mecanismo procedimental consta de la Política de Gestión de Documentos Electrónicos, que se aplica a todo el Ayuntamiento y a las entidades dependientes de él; y de una figura mediadora entre el Archivo y el Servicio de Informática, que actúa tanto en calidad de consejero como en calidad de controlador, dos de los roles típicos de la responsabilidad.

El mecanismo tecnológico está basado en un reforzamiento de la entidad funcional Ingesta, adoptada de un OAIS, que, mediante la definición de reglas de validación y servicios web, supervisa que no se transfiera al Archivo nada que no cumpla determinadas condiciones, impidiendo que aquél se haga responsable de documentos potencialmente corruptos o de cualquier otra manera inaceptables.

4. Escenario tecnológico

Para lograr este fin, el Ayuntamiento de Cartagena alcanzó un acuerdo con la compañía de software Odilo, de tal modo que el equipo de programación del primero pudiera enfocarse sobre el sistema de producción digital, mientras que el equipo de la segunda pudiera enfocarse sobre el proceso de ingesta y su pista de auditoría. El primer proyecto, si bien profundamente interconectado con el segundo, queda fuera del alcance del presente texto, por lo que no lo describiremos aquí. Por lo demás, el modelo de implantación de un OAIS será abordado por el Archivo en una segunda fase y aún no se han definido sus premisas.

Respecto al desarrollo del proceso de ingesta y pista de auditoría, el proyecto mostraba algunas complejidades, principalmente debidas al hecho de que las especificaciones técnicas mencionadas más arriba fían fuertemente sobre la interoperabilidad y las relaciones. Hasta fecha reciente, estos conceptos no formaban parte del vocabulario de los archiveros del Ayuntamiento de Cartagena, ni, para ser precisos, de los archiveros de ninguna otra administración. Por tanto, las herramientas archivísticas al uso no incorporaban respuestas a esta nueva realidad. Como consecuencia, los desarrolladores tomaron la decisión, de común acuerdo con el Archivo, de descartar las herramientas convencionales y, en lugar de ellas, utilizar tecnologías más flexibles, capaces de abordar la interoperabilidad y las relaciones:

Estas tecnologías son las siguientes:

- Arquitecturas orientadas a microservicios: “De manera breve, el estilo arquitectónico de microservicios es una aproximación para desarrollar una sola aplicación como un conjunto de pequeños servicios, cada uno de los cuales ejecuta su propio proceso y se comunica mediante mecanismos ligeros, a menudo un recurso API HTTP. Estos servicios se construyen en torno a capacidades de negocio y son desplegados de manera independiente por maquinaria de despliegue completamente automatizada. Existe un mínimo de gestión centralizada de estos servicios, que pueden escribirse en diferentes lenguajes de programación y utilizar diferentes tecnologías de almacenamiento de datos” (Fowler, 2015).
- Bases de datos orientadas a grafos: “La representación en grafos, esto es, una colección de nodos y enlaces entre nodos, da soporte a todos los aspectos del proceso de minería de datos relacional. Como una de las formas más generales de representación de datos, el grafo representa con facilidad entidades, sus atributos y sus relaciones con otras entidades ... Puesto que una entidad puede estar enlazada de manera arbitraria con otras entidades, la lógica de las bases de datos relacionales tiene dificultades para organizar los datos de manera que

den soporte eficiente y transversal a los enlaces relacionales. De manera típica, las representaciones en grafos almacenan todas las relaciones de una entidad con otra entidad. Finalmente, las bases de datos relacionales y sus representaciones lógicas no dan soporte a la visualización directa de datos y conocimiento. De hecho, la información relacional almacenada de este modo se convierte de manera típica a alguna forma de grafo para su visualización. La utilización de un grafo para representar los datos y el conocimiento extraído da soporte a una visualización directa y a una comprensibilidad incrementada del conocimiento. Por tanto, la minería de datos en grafos es una de las aproximaciones más prometedoras a la extracción de conocimiento a partir de datos relacionales” (Holder, Cook, 2007. P. 12).

- NoSQL: “Adecuado o no, el nombre intenta describir el número creciente de bases de datos no relacionales distribuidas que ha emergido durante la segunda mitad de los 2000 ... Algunos de los rasgos generales, aunque no ubicuos, que muchos de los sistemas NoSQL comparten: carecen de esquemas fijos, evitan las uniones (la operación de combinar relaciones), escalan de manera horizontal” (Lith, Mattsson, 2010. P. 16).
- Schema-less: “Una base de datos sin esquema permite que cualquier dato, estructurado con campos y estructuras individuales, quede almacenado en una base de datos. El ser sin esquema reduce la ceremonia (no se tienen que definir esquemas) e incrementa la flexibilidad (se puede almacenar todo tipo de datos sin definición previa)” (Fowler, 2013).
- Web semántica: “La Web semántica proporciona un marco común que permite que los datos sean compartidos y reutilizados más allá de los límites de aplicaciones, empresas y comunicaciones. Es un esfuerzo colaborativo liderado por el W3C con la participación de un amplio número de investigadores y socios industriales. Está basado en Resource Description Framework (RDF)” (Hawke et al., 2013).

Además, los desarrolladores propusieron el modelo conceptual de la responsabilidad de Martin Fowler para resolver el problema de las relaciones (Fowler, 1999); así como una metodología ágil basada en el Agile Methodology Manifesto (Beck et. al., 2001). Esto nos permite promover de manera fácil las interacciones entre el equipo de programación y el Archivo para, mediante sucesivas iteraciones, enriquecer progresivamente los resultados.

Este conjunto de tecnologías y metodologías viene acompañado por un conjunto de productos *open source*, especializados en la indización de logs y pistas de auditoría.

5. Resultados

Con respecto a la ingesta, el Archivo Municipal determinó como un requisito esencial el hecho de que todo expediente transferido al Archivo tenía que ser conforme con las restricciones legales derivadas del Derecho español; de lo contrario, debería ser rechazado. Esto significa que:

- Un documento es un compuesto de datos de contenido, metadatos y firma. Todos ellos deben ser empaquetados antes de ser transferidos;
- Los mecanismos de denominación de ficheros deben seguir ciertas reglas;
- Los ficheros XML que subyacen a los documentos y expedientes deben ser conformes con los distintos esquemas XSD publicados por el Ministerio de Hacienda y Administraciones Públicas;
- Los ficheros XML que subyacen a los documentos y expedientes deben satisfacer un cierto número de reglas de contenido, por ejemplo, en lo que concierne al formato de fecha o a la clasificación funcional;
- Las firmas, en el caso del Ayuntamiento de Cartagena XadES, deben chequearse, no sólo contra el esquema proporcionado por el Ministerio, sino también contra su propio esquema ETSI (European Telecommunications Standards Institute);

Estas reglas mínimas de carácter general han sido complementadas con algunas otras de carácter local:

- Los documentos deben ser inmediatamente utilizables por usuarios finales; por tanto, los ficheros XML deben venir acompañados por sus paralelos ficheros PDF, que a su vez se encuentran embebidos en base64 en aquellos XML;
- Los paquetes deben encapsularse mediante un hash, y este hash debe conservarse junto con, pero separado de, los ficheros que componen el documento;
- Para permitir que el Archivo lleve a cabo los procesos de valoración y acceso, la clasificación funcional debe asignarse de manera automática en el momento de crear un expediente;

- Los agentes deben identificarse de acuerdo con mecanismos nacionales, a saber, el Directorio DIR3, en el momento en el que actúan sobre un expediente, y se debe chequear esta circunstancia durante el proceso de ingesta.
- En el momento de la transferencia, una vez validado un paquete de información, el hash debe conservarse en el Archivo, pero enviarse también a un segundo depósito como garantía adicional de integridad.
- Una vez validado un paquete de información, debe conservarse como tal, pero también deben desempaquetarse los ficheros, a efectos de capacidad de reproducción a lo largo del tiempo.

Como indicamos más arriba, para impedir que un documento no conforme sea ingerido por el Archivo, aunque haya sido transferido por el sistema de producción, se ha desarrollado un servicio web para verificar de manera automática que todas las reglas mencionadas se han satisfecho.

Con respecto a la seguridad, en circunstancias ideales, varios depósitos, tal y como las regulaciones sugieren, no serían necesarios. No obstante, tanto para obedecer la legislación como para respetar los protocolos de seguridad del Ayuntamiento de Cartagena, hemos añadido dos depósitos al modelo conceptual OAIS. Realmente, sólo el OAIS ofrecería garantía legal, pero estos depósitos complementarios ayudan a asegurar que las propiedades de los documentos quedan salvaguardadas a lo largo del tiempo y a impedir que el documento sea repudiado en corte. Hemos llamado a esta arquitectura “meta-OAIS” (véase Figura I). El primer depósito complementario, bajo control del Archivo, permite vigilar y supervisar de manera automatizada cualquier incoherencia, error o incidente. El segundo, como hemos mencionado es resultado de un acuerdo entre el Ayuntamiento de Cartagena y la Autoridad de Certificación de la Generalitat de Valencia, que mantiene un depósito donde se transfiere también el hash de todo expediente transferido al Archivo, a efectos de chequear la integridad, en el supuesto de que exista sospecha de que ésta ha quedado comprometida.

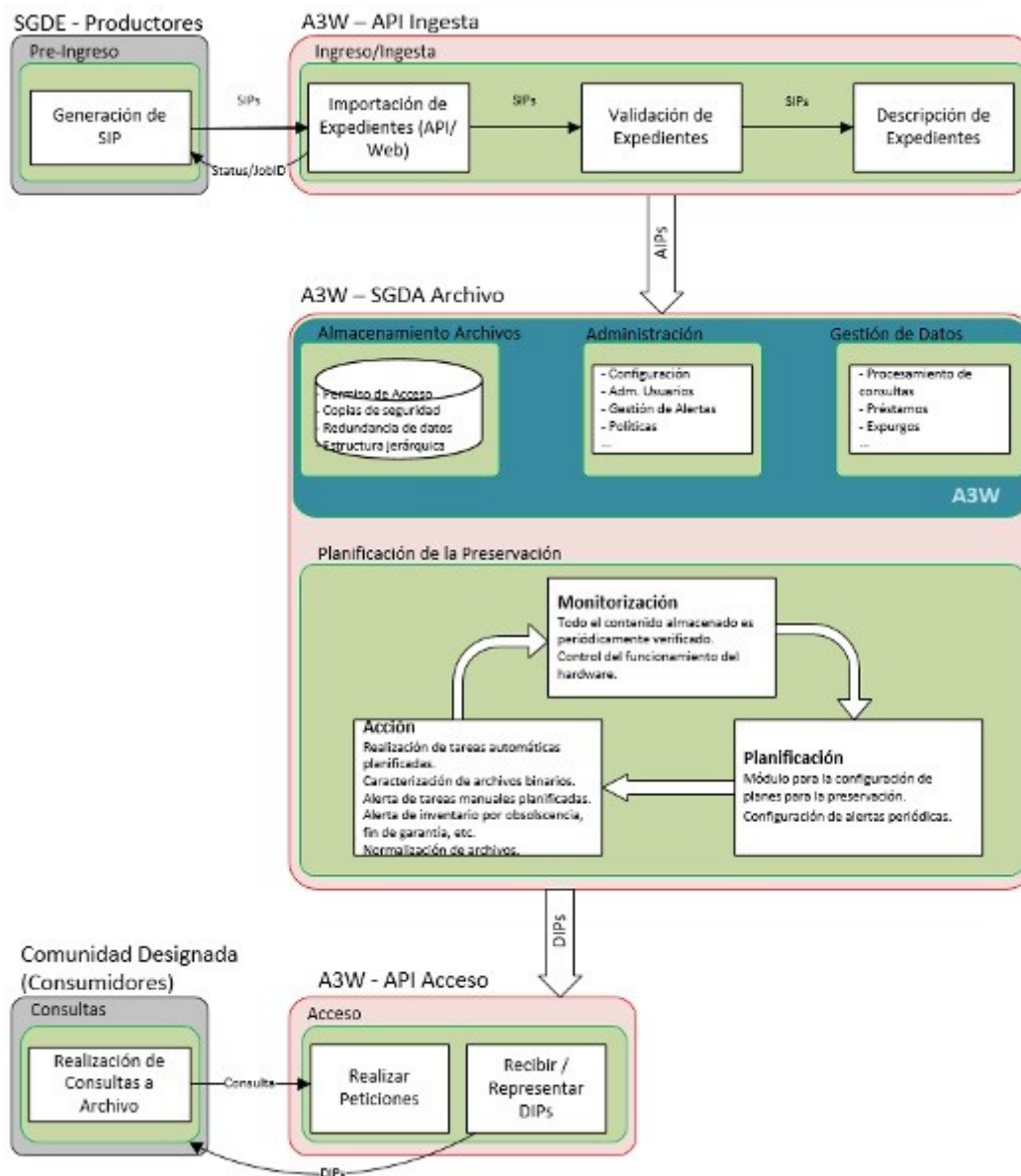


Figura I: Meta-OAIS

Agradecimientos

El autor desea agradecer las explicaciones acerca de tecnologías proporcionadas por Pablo Gutiérrez Mosquera y Ricardo den Haan Alonso. Sin su perspicacia, este texto no hubiera sido posible.

Bibliografía

Beck, Kent et al.: Manifesto for Agile Software Development. The Authors, 2001. <http://agilemanifesto.org/> [Consulta: 13 de junio de 2015]

Cartagena. Ayuntamiento: Ordenanza Municipal de Administración Electrónica. <http://www.cartagena.es/files/115-33003-DOC_FICHERO/BORM%20252.%2030-10-10.%20ORDENANZA%20MPAL%20ADMON%20ELECTRONICA.pdf> [Consulta: 13 de junio de 2015]

Cartagena. Ayuntamiento: Política de Gestión de Documentos Electrónicos. <https://seguro.cartagena.es/sedeelectronica/docs/politica_gestion_documentos_electronicos.pdf> [Consulta: 13 de junio de 2015]

Delgado Gómez, Alejandro: Anotaciones sobre el archivar en los espacios digitales contemporáneos. En: V Seminario de Arquivología, Belo Horizonte (Brasil), 4-8 noviembre 2013

Delgado Gómez, Alejandro: El centro y la equis: una introducción a la descripción archivística contemporánea. Cartagena: Ayuntamiento, 2007

Delgado Gómez, Alejandro: Documentos y poder: órdenes del discurso. En: Anales de Documentación de la Universidad de Murcia, 2010, vol. 13, pp. 117-133

Delgado Gómez, Alejandro; Cruz Mundet, José Ramón: Los archivos como construcción social. Tenerife: Anroart, 2010

España. Gobierno: Ley 11/2007, de 22 de junio, de Acceso Electrónico de los Ciudadanos a los Servicios Públicos. <<http://www.boe.es/buscar/doc.php?id=BOE-A-2007-12352>> [Consulta: 13 de junio de 2015]

España. Gobierno. Real Decreto 4/2010, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica. <<http://www.boe.es/buscar/act.php?id=BOE-A-2010-1331>> [Consulta: 13 de junio de 2015]

España. Secretaría de Estado de Administraciones Públicas: Resolución de 28 de junio de 2012, de la Secretaría de Estado de Administraciones Públicas, por la que se aprueba la Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos. <http://www.boe.es/diario_boe/txt.php?id=BOE-A-2012-10048> [Consulta: 13 de junio de 2015]

España. Secretaría de Estado para la Función Pública: Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Documento Electrónico. <http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-13169> [Consulta: 13 de junio de 2015]

España. Secretaría de Estado para la Función Pública: Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Expediente Electrónico. <http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-13170> [Consulta: 13 de junio de 2015]

España. Secretaría de Estado para la Función Pública: Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Política de Firma Electrónica y de certificados de la Administración. <http://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-13171> [Consulta: 13 de junio de 2015]

Fowler, Martin: Microservices. In: Martin Fowler. <<http://martinfowler.com/articles/microservices.html>> [Consulta: 13 de junio de 2015]

Fowler, Martin: Organization Structures. In Martin Fowler. <<http://martinfowler.com/apsupp/accountability.pdf>>. [Consulta: 13 de junio de 2015]

Fowler, Martin: Schemaless Data Structures. In: Martin Fowler. <<http://martinfowler.com/articles/schemaless/>> [Consulta: 13 de junio de 2015]

Hawke, Sandro, Herman, Ivan, Archer, Phil, Prud'hommeaux, Eric: W3C Semantic Web Activity. W3C, 2013. <<http://www.w3.org/2001/sw/>> [Consulta: 13 de junio de 2015]

Holder, Laurence B., Cook, Diane J.: Introduction. In: Mining Graph Data, edited by Diane J. Cook and Lawrence B. Holder. John Wiley & Sons, Inc., 2007.

International Organization for Standardization: ISO 14721:2012: Space data and information transfer systems -- Open archival information system (OAIS) -- Reference model. ISO, 2012.

Lith, Adam, Mattsson, Jakob: Investigating storage solutions for large data: A comparison of well performing and scalable data storage solutions for real time extraction and batch insertion of data. Department of Computer Science and Engineering, Chalmers University of Technology, 2010.

XXX: XXX