



**ANA MARÍA LÓPEZ CUADRADO
JAVIER REQUEJO ZALAMA**

Estrategias y modelos de gestión de datos archivísticos

Archival data management
strategies and models

*Ana María López Cuadrado, anamaria.lopez@cultura.gob.es
Departamento de Normalización y Coordinación
del Archivo Histórico Nacional
Javier Requejo Zalama, javier.requejo@cultura.gob.es
Archivo de la Real Chancillería de Valladolid*

*Citación: López Cuadrado, Ana María, Requejo Zalama, Javier (2021). "Estrategias y modelos de gestión de datos archivísticos". *Tábula*, n. 24, pp. 21-43*

*Recibido: 2-3-2020. Aceptado: 5-10-2020
DOI: <https://doi.org/xxx>*

Resumen analítico / Analytic summary

La selección de un modelo de datos archivísticos se ha convertido en un elemento clave para adoptar una estrategia más amplia de gestión de la información en archivos. Así, el presente artículo reflexiona sobre el último modelo de datos, basado en múltiples entidades archivísticas relacionadas entre sí, clave en las más recientes normas (NEDA en España, RiC...) y cuya implementación es posible gracias a la actual tecnología. Mediante los elementos fundamentales de dicho modelo, en este texto se analiza su viabilidad en entornos de archivos históricos, pero en archivos orientados a las fases iniciales del ciclo documental.

ESTRATEGIAS | MODELOS DE DATOS | NORMAS DE DESCRIPCIÓN | ENTIDADES ARCHIVÍSTICAS | NEDA | RiC

The selection of a archival data model has become a key element in the adoption of a broader strategy for managing information in Archives. This article reflects on the latest data model in Archives, based on multiple interrelated archival entities, key in the most recent standards (NEDA in Spain, RiC...) and whose implementation is possible due to the current technology. Using the main elements of this model, the authors analyze its viability in historical archives, but also in records management units, oriented to the first stages of the records life cycle.

STRATEGIES | DATA MODELS | DESCRIPTIVE STANDARDS | ARCHIVAL ENTITIES | NEDA | RiC

*No me hace falta ver a dónde voy...
Sólo saber de dónde vengo*
(Cars, 2006)

Cualquier sistema archivístico debe disponer de un plan o una estrategia que aborde la política de gestión de datos que se manejan para poder valorar los documentos que se han de conservar o los datos que se han de mantener. Este plan no tiene por qué materializarse en un documento al estilo de los planes estratégicos configurados en los ámbitos más gerenciales de las organizaciones, sino que puede entenderse como una serie de decisiones u objetivos que los archivos deben alcanzar con la intención de facilitar/garantizar que los usuarios encuentren y obtengan la información que buscan.

En este sentido, queremos aprovechar este espacio para lanzar y compartir una serie de reflexiones sobre estrategias, analizando no tanto el contenido detallado de los desarrollos normativos, pasados o recientes, que afectan a la gestión de datos, ni los últimos avances tecnológicos que permiten aplicar estos desarrollos en los entornos de trabajo archivísticos, sino más bien, cómo se viene aplicando la toma de decisiones o el despliegue, en un modelo de gestión de la información necesario para alcanzar los objetivos previamente identificados¹.

Y es que, a lo largo del tiempo, y dependiendo de la estrategia o del modelo que se vaya a emplear para gestionar la información que está representada en los datos que manejamos, podrán variar sustancialmente tanto las opciones disponibles como los resultados esperados e incluso el *target* de usuarios, lo cuales se pueden ver beneficiados de los esfuerzos archivísticos.

De esta forma, para analizar en qué punto se encuentra un archivo, un sistema archivístico o la archivística en general en lo que se refiere a la estrategia a aplicar, debemos tener en cuenta varios factores: el modelo de datos del que se parte inicialmente, si existe o no intercambio de datos, su grado de normalización y la forma de recuperar la información.

La descripción archivística tradicional en España ha mantenido desde sus inicios un sistema relacional muy arraigado, el cual nos ha permitido llegar hasta el punto en el que nos encontramos actualmente. Esta descripción, basada fundamentalmente en fichas catalográficas, siempre ha desarrollado este modelo relacional, pero teniendo al documento como único punto de estudio. De este modo, si diseccionamos una de estas fichas, encontramos la mayor parte de las entidades que actualmente los estándares NEDA y el RIC han comprendido como fundamentales en la archivística: el documento, su contenido, intervinientes, productor, data... Estos elementos, protagonistas en cualquier catálogo tradicional, cobran en la actualidad la misma importancia con las nuevas normas de descripción archivística por la relación que mantienen entre sí. Es por eso por lo que, nuestras bases de datos –procedentes en su mayoría del siguiente paso que se produjo en la archivística: la informatización de los ficheros– deben ser reconocidos como auténticos tesoros que, tratados con las técnicas disponibles en la actualidad, aportan un universo de datos impensable.

Siguiendo un análisis cronológico de los hechos, el siguiente hito en la descripción archivística fue la aparición de las normas del Consejo Internacional de Archivos, las cuales impusieron ciertas directrices, como la jerarquía, la no repetición de los datos, y la descripción multinivel, dejando abierta la posibilidad de relacionar las descripciones de los documentos con otras entidades ya fueran documentales o no. Esto quedó refrendado con la ISAAR CPF y la ISDIAH, y en cierto modo con la injustamente tratada ISDF, la hermana pequeña de estas normas, poco usada y menos considerada.

A la vista de estos esfuerzos normativos, se podría decir que el mundo archivístico de las últimas décadas ha estado basado en un modelo de datos jerárquico, defendido y representado por estas normas del Consejo Internacional de Archivos. Las relaciones entre los datos se centraban en una jerarquía, en la que el cuadro de clasificación y su representación mediante una especie de inventario dinámico, cobraba una especial relevancia. Incluso en muchos casos, se pretendía representar estos cuadros en los depósitos, haciendo que la estructura de la documentación correspondiera también con la estructura lógica. Por tanto, hablamos de datos, pero estructurados solo de forma jerárquica.

De cualquier modo, a pesar de tener una visión bastante clara de un modelo de datos relacional que, no cabe duda, resulta muy innovador para su época, la archivística se basaba en un sistema docucéntrico, es decir, anclado en el documento y su jerarquía, con un desarrollo de puntos de acceso entendidos inicialmente como simples elementos auxiliares de la descripción basados en literales, como mucho tesauroizados, que nos ayudarían a marcar hitos en la documentación, pero que realmente no servían de forma clara para buscar y recuperar los datos demandados por los usuarios de forma rápida y sobre todo eficiente.

En estos modelos fundamentalmente se trabajaba en local, es decir, en entornos cerrados con accesos exclusivamente desde las salas de consulta y de forma presencial. Los archivos, pese a que han tenido ambiciosos proyectos descriptivos, primero en papel utilizando las fichas, y posteriormente con las bases de datos o gestores documentales, no favorecían la consulta de tal cantidad de información fuera del entorno físico del archivo. Sin embargo, poco a poco comenzaron a surgir diferentes proyectos que pretendían solventar esta coyuntura, a través de lenguajes de intercambio de datos basados en XML. Los esquemas EAD, EAC-CPF y EAG nacieron con esa visión, la de ser un motor para intercambiar información archivística a través de Internet, pero siguiendo esta misma pauta jerárquica que plasmaban las normas internacionales de descripción. De hecho, estas normas y sus respectivos esquemas en XML se autoalimentan entre sí, con lo que el modelo de datos seguía siendo el mismo, aunque en un entorno más abierto y accesible.

Con respecto a la recuperación de la información, los sistemas de datos archivísticos se basaron en un primer momento en la construcción de índices o puntos de acceso que nos permitieran recuperar datos, aunque inicialmente de forma muy subsidiaria. Ciertamente es que con el paso del tiempo se avanzó un buen camino trabajando en tesauros, de nuevo un sistema jerárquico de representación de datos. Pero la búsqueda por texto libre ganaría la partida imponiéndose entre las estrategias de gestión, pasando a ser la forma más usada para recuperar la información. Con las búsquedas por texto libre se podía recuperar cualquier dato en cualquier punto de las descripciones, algo que resultaba totalmente novedoso e impactante. Las búsquedas tipo *Google* con una única barra de búsqueda pasarían a estandarizarse y convertirse en la piedra angular de los sistemas de gestión documental. Pero esta forma de entender la recuperación de la información se encontró con un problema: el ruido. La masa de datos empezaba a ser cada vez más grande y los motores de búsqueda no podían recuperar tantos datos en búsquedas sencillas. Los resultados ofrecían mucho ruido informativo, o incluso cada vez más frecuentemente, se obtenían respuestas que tenían este aspecto: “*Por favor, acote la búsqueda porque hay muchos resultados*”.

Es en este contexto donde se ha de incluir la creación de la Comisión de Normas Españolas de Descripción Archivística y la adopción de los estándares NEDA. Se trataba de imprimir un cambio de mentalidad a través de un esfuerzo de abstracción y una especie de *tabla rasa* que debiera hacer la comunidad profesional archivística, como replantear sus conocimientos y valores presuntamente tradicionales sobre la gestión de la información y reclamar la necesidad de interactuar y cooperar intensamente con otros ámbitos profesionales como es el de las tecnologías de la información. A partir de este momento, las redes neuronales que Internet va difundiendo, el *Data Science* que ya se ha empezado a cristalizar en otros sectores y la necesaria interdisciplinariedad para alcanzar el cumplimiento de los objetivos, empiezan a tener sentido también en los archivos.

Como ya hemos comentado en párrafos anteriores, el modelo de datos planteado a través de los estándares NEDA viene avalado a nivel internacional por el desarrollo similar que se plantea a través del RIC (aún sin publicar la versión definitiva), basándose ambos en la relación entre diferentes tipos de entidades y sus atributos. Cierto es que la premisa no resulta radical ni rupturista, pero sí es novedosa e impactante en cuanto al análisis y la estructura de los datos, de cara a facilitar nuevas formas de visualización o recuperación de la información. Ya no solo se estudia el documento con su serie o su fondo como relaciones de contexto, sino que se parte de las entidades documentales (el documento, la serie, el fondo...) para analizar y dar a conocer también entidades no documentales, por ejemplo, a su agente productor, a los conceptos relacionados, a la función que testimonia la serie o al lugar en el que el documento se data. Incluso se puede ofrecer acceso a información adicional sobre el soporte, el tipo documental, las relaciones con otros documentos, los instrumentos de descripción o la bibliografía, todo ello a través de relaciones que generan mucho más conocimiento y amplían nuestras redes neuronales de información, lo que al final revertirá en un mayor conocimiento para nuestros usuarios.

Es decir, el nuevo paradigma planteado en los estándares NEDA o en el futuro RIC defiende que es necesario efectuar un estudio pormenorizado de los datos a priori, para mejorar la recuperación de la información a posteriori. En este caso, las entidades no documentales no son meros índices, literales, ni puntos de acceso, sino que cobran importancia por sí mismas dentro del modelo de datos, y realmente sirven para mejorar de forma notable la recuperación y la satisfacción de los usuarios ante la demanda de información. Ya no nos encontramos ante un modelo docucéntrico, sino que es un modelo en el que el documento es la raíz, pero las entidades que lo rodean cobran la importancia necesaria para enriquecerlo, estudiarlo y, sobre todo, hacerlo más visible y recuperable desde múltiples puntos. Y si a este sistema le añadiésemos una normalización exhaustiva de los datos que se vienen manejando, entonces el intercambio de información resultaría incluso más rápido y fácil, dependiendo tan solo de los recursos tecnológicos disponibles que lo faciliten, ya que, desde el punto de vista archivístico “*informativa*”, este resultaría totalmente plausible per se.

Debido a los avances que ha supuesto para la archivística este nuevo modelo, en la actualidad muchos de los grandes hitos previos en materia de descripción archivística, como la jerarquía de las descripciones o la presentación de estas mediante inventarios dinámicos que asemejan cuadros de clasificación, empiezan a tambalearse porque ahora ya no hay una necesidad acuciante de crear esa jerarquía. Este tipo de representación gráfica de la información ya no resultaría tan necesario si comenzamos a representar esa misma jerarquía a través de relaciones estructuradas, mediante otros mecanismos (por ejemplo, vínculos informáticos entre descripciones).

En esta misma línea, otro pilar que también se tambalea en la aplicación de esta nueva estrategia es el concepto de “*descripción multinivel*”, asentado a través de las normas del Consejo Internacional de Archivos, y que empieza a ser sustituido por el de “*descripción multientidad*”, que se convierte en el elemento clave en este nuevo modelo de datos.

Y algo similar ocurre con los puntos de acceso entendidos como simples índices o literales. En los nuevos modelos de datos estos pasan a carecer totalmente de sentido, perdiendo la partida frente a las entidades no documentales que pueden ser objeto de descripción por sí mismas. No sirve de nada sacar una o varias palabras (un nombre, un lugar, etc.) como un índice en un documento cuando, a la hora de utilizar en el futuro ese índice, no va a recuperarse nueva información y tampoco va a llevar a un resultado unívoco. Por ejemplo, ¿de qué sirven los índices genéricos referidos a nombres de persona que se utilizan como una bolsa de múltiples personas con el mismo nombre y apellidos, incluso de diferentes épocas históricas y sin posibilidad de discriminar entre ellos, si no es accediendo a cada una de las descripciones de los documentos en los que están indizados? Posiblemente, al principio o en sistemas que manejan un volumen de datos muy limitado, estas prácticas podrían llegar a tener cierta utilidad ante ciertos trabajos de organización o ante las demandas de un cierto perfil muy concreto de usuarios. Pero siendo realistas, en entornos en los que ya existen búsquedas por texto libre de forma consolidada, los índices genéricos de este tipo “*ofrecen más ruido que nueces*”.

Frente a todo esto que venimos mencionando, resulta comprensible que surjan preguntas o reflexiones acerca de lo que esta nueva estrategia conlleva ¿Realmente sabemos lo que significa un modelo de “*múltiples entidades*” para los Archivos? ¿Estamos dispuestos a prescindir de la presentación y visualización del inventario dinámico en pro de un análisis más minucioso de nuestros fondos y de otras formas de visualización de los datos? ¿Estamos capacitados para deconstruir nuestros esquemas para poder representar la misma información desde otro punto de vista, pensando en el usuario y sus formas de plantearse las búsquedas de la información? ¿Realmente queremos prescindir de los puntos de acceso tradicionales para introducirnos en el proceloso mundo de las entidades no documentales? Y en última instancia, ¿en los archivos realmente se quiere aplicar esta nueva forma de entender la archivística? ¿O es que en realidad no es posible llevarlo a cabo y todo esto no deja de ser simplemente un “*brindis al sol*”?

Ante todas estas preguntas, vamos a intentar hacer una revisión de la situación actual, y para ello vamos a separar nuestro análisis en dos ámbitos: Primero los archivos “*históricos*”, y luego los archivos “*administrativos*”. El porqué de esta separación viene provocado, no porque entendamos que exista una dicotomía profesional entre ambos entornos sino porque, para conseguir sus fines, aplican esta estrategia a través de soluciones diferentes para así alcanzar sus respectivos resultados.

El nuevo modelo de datos en los archivos “históricos”

Los archivos históricos tienen entre sus funciones principales la conservación y la difusión del patrimonio documental que custodian. Para ello, el control de la documentación resulta fundamental, entendiéndose aquí la unidad documental física como la clave para el buen funcionamiento del servicio, para su consulta, su protección, su descripción, el movimiento de fondos... Es por ello por lo que tiende a asegurarse que, en los archivos históricos, la signatura es la pieza fundamental en los trabajos archivísticos (por encima, incluso del resto de elementos informativos).

Si nos centramos en los procesos de identificación y clasificación de los fondos archivísticos históricos, encontraremos diferentes tipos de exhaustividad en las descripciones. Dependiendo de los medios personales, de los recursos disponibles y de la documentación que manejemos, se apuesta por una descripción a nivel de documento, a nivel de expediente, de serie... O incluso a nivel de unidad de instalación (legajo, caja, carpeta...), ya que este último tipo de elemento permite controlar la documentación que se encuentra en los archivos con una gran cantidad de fondos, permitiendo conocer el contenido de esas unidades, facilitando la realización de inventarios topográficos y en última instancia, ofreciendo un mejor servicio a los usuarios.

En los archivos históricos hay ocasiones en las que es obligado alcanzar un nivel de información muy exhaustivo, debido a la posesión de ciertos elementos excepcionales o característicos en las entidades documentales, a la necesidad de relacionarlas con su imagen, a la importancia eventual que pueda adquirir un documento, etc. Así, mapas, planos, dibujos, fotografías, pergaminos, películas, o documentos de otra naturaleza, pueden precisar de una descripción más exhaustiva que el resto de los documentos del archivo. En cambio, hay ocasiones en las que prima la necesidad de conocer qué se custodia, sin necesidad de entrar en detalles tan precisos, para poder ofrecer al usuario un servicio básico de información y consulta.

Poniendo el foco en la descripción de los documentos, y teniendo en cuenta que las tareas de identificación, clasificación y ordenación se hayan realizado correctamente, se podría afirmar que el nuevo modelo de datos ya se viene aplicando en los archivos históricos, aunque sea de forma parcial. De manera general, se entiende todo lo que significa y supone el modelo multientidad, y al menos se aplica todo aquello que ya venía definido en las normas del Consejo Internacional de Archivos: las relaciones de jerarquía, la relación con el productor y el uso de puntos de acceso. Ahora bien, el resto de las relaciones que se contemplan en el modelo de datos no se suelen desarrollar, aunque como en todo, hay excepciones.

El primer problema para la implementación de este nuevo modelo de datos es que los sistemas de información actuales no están preparados para asumir estos cambios, ni han sido diseñados para permitir una evolución de estas características. Aunque se pueden encontrar instrumentos, aplicaciones informáticas, herramientas tecnológicas o ejemplos reales que facilitan la adopción de estas novedades, todavía cuesta encontrar modelos prácticos que realmente permitan implementar un modelo completo de descripción archivística, en el que todas las entidades, documentales y no documentales, se relacionen entre sí.

El segundo problema es el grado de exhaustividad que pueden precisar las descripciones de este tipo, no tanto por el volumen de documentos o la cantidad de información sino por el volumen de entidades que hay que identificar y describir de forma individualizada, para luego poder interconectarlas. De un modo más concreto, siguiendo los criterios de los estándares NEDA, si quisiéramos describir todas las entidades no documentales que pueden aparecer en los fondos de un archivo histórico, habría que identificarlas y describirlas adecuadamente, no solo hacer meros literales o representarlas sin más a través de un nombre. Tendríamos que identificar cada una de las entidades de forma unívoca y clara, de tal manera que cuando aparezca en otro documento, serie o fondo, podamos vincularla sin dudas con esa nueva entidad documental. Y ese trabajo no es rápido ni sencillo, como puede serlo crear un simple índice o redactar un campo de texto libre.

Esta nueva forma de descripción hace que el proceso descriptivo se ralentice, o más bien que los tiempos necesarios para describir un documento varíen sustancialmente frente a lo que se venía haciendo anteriormente puesto que, ahora, habría que dedicar más tiempo a describir entidades no documentales, estableciendo continuas relaciones con descripciones creadas previamente. Todo esto lleva a plantear hasta qué punto es rentable describir de esta nueva forma, cuando los archivos históricos conservan todavía mucha documentación sin ni siquiera identificar.

Hay que reconocer que esta forma de trabajo ha de ser vista a largo plazo y saber que en los momentos en los que se empiezan a dar los primeros pasos dentro de estas estrategias, todo está por hacer. Por ello, si en nuestro sistema queremos describir entidades no documentales y no tenemos todavía ninguna descrita o accesible desde otro lugar, cualquier entidad no documental que aparezca asociada a la documentación, tendría que ser estudiada, analizada y si fuera preciso, creada en nuestro sistema antes de poder establecer la relación con el documento. Pero, con el tiempo, cada vez habrá más entidades no documentales disponibles y los trabajos se podrían convertir en algo más rápido y provechoso, ya que las relaciones podrán crecer de forma exponencial, permitiendo a los usuarios encontrar más información de la que a priori ellos no conocían.

Con todo esto no pretendemos dar validez a posibles críticas acerca de la escasa aplicabilidad o falta de realismo del nuevo modelo de datos en los archivos históricos. Muy al contrario, a medio plazo, si se sigue describiendo, basándose en esta estrategia, los resultados serán inconcebiblemente mejores para los usuarios, los cuales la mayor parte de las veces permanecen ajenos a los problemas internos que puede conllevar esta forma de trabajo, y lo que desean es satisfacer su necesidad de información. Que un usuario pueda navegar entre descripciones, ya sea de documentos, conceptos, personas, lugares, materias, soportes, instrumentos de descripción, fuentes bibliográficas, etc. sin saber quién ha producido toda esa información o incluso sin saber en qué archivo se encuentran físicamente los documentos que le puedan interesar, es un avance impensable hace tan sólo un par de décadas.

Tampoco queremos decir que los archivos no tengan que seguir manteniendo procesos de identificación o descripción de los documentos, tal y como se venían haciendo, a niveles de macro descripción. Consideramos que el conocer los órganos productores de la documentación y la historia archivística de los fondos permite orientar de la mejor forma posible a los usuarios, para que éstos puedan resolver las pesquisas o investigaciones en las que puedan estar centrados y de las que, en última instancia, también podemos aprender en los archivos.

Pero el usuario no tiene que ser un experto en órganos productores de la documentación para poder obtener información. Los expertos en ello debemos ser los archiveros, pero nuestra misión debe ser ofrecer el mayor número de herramientas y mecanismos posibles para que los usuarios encuentren lo que desean y lleguen a sus propias conclusiones, sin tener que precisar de un conocimiento detallado de la entidad que fue la productora de los fondos, de las funciones que dicha institución desempeñaba, de la legislación que la creó... Para eso estaremos nosotros, los archiveros, con competencias para identificar debidamente esta información, apoyarnos en herramientas que nos permitan desglosar todos estos datos en elementos y relacionarlos en todos los sentidos, definiendo del modo más preciso la relación entre las diversas entidades.

Los estándares NEDA y el futuro RIC asientan el concepto de entidad, pero donde han plasmado la mayor diferencia y el punto de inflexión con los estudios y normas preexistentes ha sido en dos puntos. Por un lado la definición de las relaciones, planteando nuevos conceptos, como la realidad por la cual dos entidades se pueden relacionar en varias ocasiones e incluso por motivos diferentes y esto debe poder ser recogido en nuestro sistema de gestión (por ejemplo, Felipe II se puede relacionar con Carlos I por parentesco, pero también se puede relacionar por ser, el primero, sucesor del segundo en el trono); por otro, la importancia imprimida en las funciones que son, como hemos indicado anteriormente, las entidades archivísticas menos desarrolladas.

Para todo ello, el archivero ha de conocer en profundidad sus fondos, las instituciones productoras y remisoras de estos, definir las entidades que participan

y estudiarlas. Porque de ahí, de toda esta historia archivística y ese contexto histórico, nace el conocimiento necesario para construir nuestros sistemas de gestión y su adecuación a nuestros archivos particulares y sus necesidades. Incluso se coge la perspectiva necesaria para poder disociar los datos y reconstruirlos posteriormente de múltiples formas según nuestra necesidad, o la del usuario que demanda la información.

Este análisis contextual, si queda plasmado en nuestros instrumentos de trabajo (es decir, en nuestros sistemas de gestión documental) no solo como elementos incorporados a nuestras descripciones, sino como entidades individualizadas relacionables entre sí, nos dará como resultado una estructura interconectada, que podrá ser representada de múltiples formas a través de las herramientas informáticas y del *Data Science* disponible. Se podrán crear hasta múltiples cuadros de fondos, inventarios funcionales, orgánicos, mixtos... No tendremos por qué elegir una sola forma de representar las descripciones y podríamos “jugar y trastear”, eligiendo una versión distinta en cada momento para visualizar los datos que quisiéramos, ya que la información estará debidamente desglosada, trabajada y estructurada. Todo lo demás sería cuestión de dejarlo en manos de la informática aplicada a nuestro trabajo.

Esto es el gran beneficio que puede aportar este modelo de datos a los archivos históricos. Y aunque todavía no se ha llegado a desarrollar plenamente, hay ejemplos de archivos e instituciones, que están intentando gestionar sus fondos basándose en estas pautas, planteándose sistemas de gestión que tratan de aportar la flexibilidad que necesita este tipo de estrategias y descripciones. Es el caso por ejemplo de los Archivos Estatales con PARES, aunque también hay comunidades autónomas como Galicia, Andalucía, Aragón o Navarra que van en una línea similar. E incluso archivos de instituciones más concretas, como puede ser el Archivo General de Palacio o el Archivo de la Real Academia de la Lengua, que pretenden introducir en sus prácticas archivísticas esta estrategia. Aun así, no es fácil encontrar instituciones que opten por abordar esta filosofía de gestión en su totalidad para todos sus fondos y todavía es necesario solventar muchos hándicaps al respecto.

El nuevo modelo de datos en los archivos “administrativos”

Pero los archivos históricos no son los únicos que ofrecen un buen encaje en esta estrategia. Y es que los archivos que podríamos llamar “administrativos”, denominados así por venir desempeñando su labor en las fases iniciales o intermedias del ciclo vital de los documentos, tienen una capacidad de adaptación similar o incluso mejor a este modelo de datos, aunque de forma muy diferente.

Para empezar, en los archivos administrativos el intercambio de información no se plantea como un fin o un resultado por sí mismo, sino como una necesidad

obligada por las diversas unidades a las que dan servicio de apoyo y soporte. Si en los archivos históricos el intercambio de información podía entenderse como el posible resultado de un objetivo cumplido, en el entorno administrativo, la interoperabilidad de la información se plantea como algo de mucha mayor relevancia, sobre todo en la fase de identificación y valoración, previa a la creación de los documentos.

Otra diferencia con respecto a la forma de abordar esta estrategia en los archivos administrativos es que no se busca tanto el detalle informativo que pueda extraerse de la documentación, a diferencia de lo comentado anteriormente para determinados casos que pueden plantearse en los archivos históricos, sobre ciertos documentos particulares o sobre ciertas agrupaciones documentales. Esta falta de exhaustividad informativa en los archivos administrativos, básicamente se produce porque la información gestionada es cada vez mayor y además tiende a repetirse. Por ello, lo que se busca es una mayor eficiencia en su manejo y gestión. De ahí, la importancia que se da en estos entornos administrativos al manejo y la calidad de los metadatos gestionados, no tanto por la información excepcional que puedan contener, sino de cara a planificar otros procesos de gestión documental como la valoración.

Otra característica clave en la implementación de esta estrategia en este tipo de archivos es que el acercamiento a la gestión de los datos no se hace exclusivamente desde la unidad de archivo, como se venía planteando en el antiguo modelo de datos o como se hace en los archivos históricos. En los entornos administrativos, los datos son gestionados por unidades archivísticas, en colaboración con otras unidades o servicios en los que los archivos no tienen por qué estar presentes (aunque quizá debieran). Por ejemplo, la gestión de los catálogos de servicios y actividades, los repositorios de procedimientos o los directorios institucionales, suelen ser gestionados y mantenidos por unidades ajenas al archivo, pese a ser considerados piezas archivísticas clave según el nuevo modelo de datos. Sí es cierto que existen casos en los que archivos y archiveros, ya sea por asignación directa competencial o por interés propio, participan activamente en su gestión y mantenimiento, o incluso los gestionan directamente, pero lo más común es que la gestión de esos datos se comparta o delegue.

Y es que precisamente porque los archivos administrativos han de ser mucho más rápidos y eficientes a la hora de gestionar los datos, y porque la colaboración entre unidades es mucho mayor, otra característica fundamental en la implementación de esta estrategia, dentro de los entornos administrativos, es que en ellos se opta por un mayor pragmatismo en general.

La anticipación de los procesos de gestión documental, dentro del ciclo vital de los documentos, a fases previas a la creación de los propios datos, no se fundamenta en el hecho de que se desee dar un mayor protagonismo a los archiveros dentro de sus organizaciones, por una cuestión de aspiraciones profesionales.

Es más, la necesidad de disponer de un sistema de gestión de la información basado en múltiples entidades de acuerdo con la estrategia que venimos planteando, no se hace basándose en la premisa de que, con una mayor inversión económica, toda la problemática quedaría solucionada. En realidad, la defensa de esta estrategia se basa en el hecho de que actualmente las propias organizaciones (instituciones públicas o privadas, con sus respectivas funciones, normas y competencias) están concebidas de forma multientidad y desde un punto de práctico, resulta mucho más beneficioso a medio y largo plazo, tanto para los archivos como para sus usuarios, mantener esa misma estrategia en general.

Conclusiones

No podemos analizar el escenario en el que nos encontramos archivísticamente hablando sin dejar de lado el contexto en el que nos movemos. El estudio de los datos y la web semántica ha evolucionado exponencialmente, junto al desarrollo del *data science*, lo que nos ha permitido poder llevar de la mano el nuevo pensamiento lógico abstracto archivístico hacia la creación de nuevos sistemas de gestión de documental que pudieran plasmar esta realidad. Así, han comenzado a nacer sistemas que buscan dar soluciones al nuevo modelo de datos, permitiendo una mayor normalización con la que conseguir una mejor recuperación de la información.

Igualmente se está consiguiendo una mejora en el intercambio de datos mediante dos vías, por un lado, avanzando y renovando los lenguajes de intercambio hacia esta nueva forma de entender la archivística y, por otro, aprovechándonos de las cosas que la web semántica nos permite hacer. Pero en esto todavía estamos “en pañales”. Necesitamos que haya más ingenieros de datos que “descubran” los recursos que tenemos en los archivos y se vuelquen con nosotros explotando la información que custodiamos. Necesitamos más proyectos de investigación que nos permitan estudiar las posibilidades que la ciencia de la información nos brinda, pero que sin datos no pueden desarrollar, poner en marcha, probar... Conjugación ciencia y archivística es una tarea que debe ir de la mano y no quedar reducida al simple diseño o selección de una herramienta para describir. Debemos mostrar y demostrar que los archivos son nuestra memoria y que la sociedad necesita de ellos, pero para ello necesitamos de otros expertos que nos ayuden a conseguirlo.

Dejando a un lado la parte tecnológica o científica de esta nueva forma de entender la archivística, con estos párrafos hemos pretendido dejar claro que este nuevo modelo no está tan alejado del modelo tradicional. No se ha inventado la rueda, pero sí se ha perfeccionado para que pueda ser más rápida, más eficiente, mejor...

El cómo tratar la documentación es perfectamente conocido por los archiveros, pero no tanto analizar de forma individualizada cada dato identificado, e intentar representar todo ello con herramientas que permitan relaciones múltiples y multidireccionales, capaces de definir y parametrizar esas relaciones y que estas sean escalables. Los objetivos por conseguir deberían ser intentar no trabajar con esquemas de información apriorísticos y dejar que el propio estudio de la documentación y sus relaciones, teplamen la estructura desde diferentes puntos de vista, sin que ninguno de ellos tenga que prevalecer sobre otro. También pretender representar la información a través de entidades documentales y no documentales, sin perder de vista la raíz, el documento, pero sin encumbrar a este, olvidando a todos sus hermanos.

Otra tarea pendiente sería conseguir una representación de la información movible según las necesidades del usuario, permitiendo una fácil recuperación adaptable a las necesidades de cada persona. Y, como último punto, lograr un intercambio de información entre archivos, pero yendo más allá de las entidades culturales o de la administración, dependiendo del tipo de archivo y las necesidades que tengamos.

Perseguir un intercambio real de la información es ahora mucho más plausible si conseguimos, por un lado, asumir como estrategia básica de gestión este nuevo modelo y por otro lograr la normalización y la simbiosis con el *data science*, la web semántica y los avances tecnológicos –con la finalidad de conseguir que todos estos últimos sean útiles para nuestro campo.

Pero estamos en ello, el “*archivero de a pie*” sigue trabajando para mejorar, aunque pueda ser reticente al cambio. Hoy en día hay muchos estudios que dicen que salir de la zona de confort no tiene por qué ser necesario siempre. A veces, mantenerse en la zona de confort es bueno... y en la archivística, no debemos tirar por la borda lo hecho hasta ahora. El trabajo realizado por nuestros predecesores ha sido y es maravilloso, y si se hizo de esa forma, fue por alguna razón, partiendo de una realidad concreta y con unos recursos concretos, que debemos conocer para poder apreciar aún más los esfuerzos realizados y los resultados obtenidos. Debemos aprovechar y reconocer los trabajos previos, sin desdeñar la nueva forma de entender la archivística. Conjugar la tradición con el nuevo modelo de datos y las demandas de una sociedad digital es la meta que la comunidad archivística debería fijarse a medio plazo. No partimos de cero, debemos apreciar el trabajo hecho, pero hay que sacarle el mayor partido posible, aprovechándonos de la ciencia y la tecnología disponible, para conseguir no solo encumbrar lo hecho hasta ahora, sino afrontar los desafíos que se vayan planteando.

Notas

¹ Nota de los autores: La ponencia ofrecida en el VIII Congreso de Archivos de Castilla y León, así como el presente texto, han sido concebidos como una reflexión general sobre el estado del arte relativo a las normas de descripción, estrategias y modelos de datos. Por ello, hemos optado por no incluir referencias a las normas citadas sobre las que se lanzan las reflexiones, así como a los posibles casos o modelos que se puedan llegar a citar como ejemplo, puesto que todos ellos son fácilmente localizables en Internet.