

## JORNADA SOBRE LA GESTIÓN DOCUMENTAL EN LA ERA DE LOS DATOS

ALDEE empezó con fuerza el año 2023 celebrando el 26 de enero una **jornada formativa** en el ámbito de la gestión documental en las dependencias del Gobierno Vasco. Este acto fue fruto de una colaboración estrecha entre el **Departamento de Gobernanza Pública y Autogobierno** y la propia asociación, siendo su objetivo resolver uno de los problemas incipientes al que se enfrentan las organizaciones: **la gestión a corto y largo plazo de los datos** recogidos o creados por la misma en el ejercicio de sus funciones.



No hay mejor comienzo para un evento de este tipo que un alegato en favor del **trabajo colaborativo** y de la importancia del **asociacionismo profesional**. Reproduciendo las palabras de **Fernando Juárez**, presidente de ALDEE: *“Las administraciones y la sociedad necesitamos asociaciones fuertes. Solos cada uno por nuestro camino, sin hablar con el vecino de al lado, no vamos a conseguir nada”*. La jornada en sí misma daba la razón a Fernando, pues la idea nació de una necesidad específica de los profesionales que trabajan en el Archivo General del Gobierno Vasco y que la asociación tomó como suya para compartirla con el resto de profesionales del sector.

Solamente con la lectura del nombre del evento se presagiaba un contenido novedoso y de gran complejidad: **“La gestión documental en la era de los datos. Valoración y preservación digital de datasets”**. El tema y los ponentes elegidos habían creado una gran expectación, casi 90 personas inscritas en la modalidad presencial y alrededor de 400 en la online. Acudirían profesionales de distintos sectores (archiveros, bibliotecarios, informáticos, ...) y de todos los lugares de la geografía española e iberoamericana (en modalidad online en su mayoría, evidentemente). ¿Se cubrirían estas expectativas tan altas?

En la ponencia inaugural **Jordi Serra** nos hizo sentir incómodos desde el principio, indicando que la gestión documental y la gestión de datos eran las dos caras de una misma moneda. Y nos invitaba a salir de nuestra **zona de confort**, esa zona en la que pensamos que la gestión de datos queda fuera de nuestra competencia. Replicando sus palabras: *“...pero ¿qué está sucediendo ahora?, que toda la información está en forma de datos, incluyendo los documentos. Los documentos también son explotables en forma de datos”*. Jordi nos hizo mirar al futuro, donde el **modelo de gestión documental** debía evolucionar para transformarse en otro modelo que



diera cabida a los datos. Todavía retumba en nuestros oídos la siguiente frase: *“Los sistemas de gestión de los datos son clave, son las nuevas series documentales del siglo XXI”*. Sin tiempo para recuperarnos del susto, siguió implacable con sus palabras: *“Lo que realmente se está llevando el gato al agua es la calidad del servicio que está recibiendo el usuario final, que es el que tiene que consumir los datos y tomar decisiones con estos datos, y que lo hace a través*

*de unas herramientas que han cambiado totalmente”*. Ese nuevo modelo propuesto por Jordi no es otro que el **modelo proveniente de la gobernanza del dato**, del que nos explicó sus pilares fundamentales (provenientes del estándar internacional **ISO-24143**), donde la gestión documental también tiene un rol importante. En primer lugar, se debe incluir la **datificación en origen** (junto con la normalización). Después, debe hacerse una **gestión lo más automatizada** posible de los datos. En tercer lugar, esos datos deben acabar en un nuevo **archivo electrónico datificado** (nueva orientación al dato del archivo). Y, por último, el ponente acabó proponiendo que la **explotación y analítica de los datos** contenidos en ese nuevo archivo electrónico pasaran a formar parte de la competencia del archivo y de la gestión documental. Al final de su ponencia nos emplazó a la tarde, donde nos enseñaría un ejemplo práctico de este nuevo modelo.

**Pilar Campos** comenzó su ponencia siguiendo el hilo argumental de Jordi, reafirmando el aterrizaje de la era de los datos: *“No podemos decir que los datos son simplemente humo. Nos aplica en nuestras organizaciones y nos aplica a nivel legislativo. A nivel organizativo han aparecido las Oficinas del Dato en múltiples instituciones, cuyas funciones acaban pareciéndose mucho a algunas de las funciones del archivo: buscar la autenticidad y la calidad de los datos, ser los custodios de los datos. Así que, aliados o nosotros. Id con cuidado, que no se nos coman la tostada. Es importante saber qué funciones tenemos nosotros con esos datos y no mirar hacia otra parte”*. Después, su presentación giró hacia una explicación más didáctica, definiendo los **conceptos fundamentales** que se iban a tratar durante toda la jornada: data item (dato), data (datos), datos estructurados y datos no estructurados, **bases de datos**, sistemas de gestión de base de datos, **datasets**, documentos y data record. Nos presentó los retos a los que nos vamos a enfrentar, nos propuso soluciones concretas e identificó problemas asociados a la preservación de los conjuntos de datos (datasets) y de las bases de datos. Y sin hacer pausa dramática nos respondió afirmativamente a las dos cuestiones principales que revoloteaban en el ambiente: *“¿Tendremos datasets en los archivos? Sí. ¿Nos pedirán archivar bases de datos? Sí.”*. Al igual que Jordi, la ponente nos hacía reflexionar sobre el concepto de **serie documental**: *“¿Igual el concepto de serie nos lo tendríamos que empezar a replantear?”*. Pilar hizo hincapié en uno de los problema principales de la preservación: los datasets pueden contener datos de diferentes



series documentales. En cuanto a la preservación de bases de datos nos emplazó a profundizar en el **proyecto e-ARK** (en su inicio proyecto **SIARD**) y su **Database Preservation Toolkit**.

Partiendo en dos la ponencia de Pilar, se produjo el merecido **descanso**. En él tuvo lugar un espacio para el encuentro entre los profesionales que asistieron presencialmente a la jornada. Esta fue la parte que no quedó grabada para la posteridad, sin embargo, fue una oportunidad única para debatir ideas, conocer a otros profesionales y hacer contactos para el futuro.

**Irene Zumalde** explicó los retos y problemas a los que se enfrenta el **Sistema de Archivo del Sector Público de la Comunidad Autónoma de Euskadi** en el ámbito de los datos. Poniendo el punto de partida en la reciente *Ley 5/2022 de Gestión Documental Integral y Patrimonio Documental de la Comunidad Autónoma del País Vasco*, la ponente fue explicando el complicado entramado que describe la ley y su sistema centralizado. A continuación, enumeró los escenarios a los que se enfrenta el sistema, enumerando los **tipos de aplicaciones** que generan datos en este ámbito y que todavía no pasan por el archivo: aplicaciones de gestión interna que no generan documentos, aplicaciones de tramitación que generan datos básicos de la misma, aplicaciones complejas de negocio, bases de datos con fines estadísticos, aplicaciones de contabilidad, etc... De cada tipo puso ejemplos reales que había en su sistema de archivo, dando así credibilidad a toda su exposición. Su propuesta pasaba por dar una solución particularizada a cada procedimiento, aunque admitió que todavía ni existían esas soluciones ni estaban operativas. Hizo hincapié en que para superar los retos de las administraciones públicas alrededor de la gestión de datos era imprescindible la **cooperación interdisciplinar**: *“Es evidente*

*que no es una tarea solo de los gestores documentales Es un proceso que debe asumir la entidad propietaria de estos nuevos documentos, en un trabajo colaborativo interdisciplinar entre productores, informáticos y gestores documentales”.*

Para acabar la jornada matutina **Beatriz Franco** explicó el trabajo que viene realizando *el Consejo de Archivos de la Comunidad de Madrid* en sus diferentes mesas de trabajo. En concreto la labor del **Grupo de Trabajo para el Tratamiento de Bases de Datos** perteneciente a una de sus mesas, la *Mesa de trabajo de archivo y documento electrónico*. La razón de la creación de este grupo de trabajo se basa en que en nuestra legislación solo se está hablando de documentos y expedientes electrónicos y que no se habla en ningún momento de bases de datos. Sin embargo, la realidad es que las administraciones están gestionando y alimentando todo tipo de bases de datos, alineándose con las tesis de Irune Zumalde. El objetivo del grupo de trabajo es crear un documento donde se explique una **metodología para la valoración de la información contenida en bases de datos**. La ponente abogó porque estos grupos fueran los más interdisciplinarios posibles para que no se pierda ningún punto de vista. Para facilitar ese trabajo de valoración, el documento, que está aún en fase borrador, propone **dos cuestionarios**: uno que sirve de guía para determinar si la información contenida se debe conservar o no (lista de verificación de criterios justificativos del valor secundario) y otro cuestionario con una lista de comprobación para validar la calidad de los datos y las extracciones de información. Los anexos del documento son de gran interés porque recogen las experiencias ya existentes en valoración de la información contenida en bases de datos, más las referencias normativas y bibliográficas que ha recopilado el grupo de trabajo. Alineándose con todos los ponentes, una de sus conclusiones fue: *“Claramente nuestra mirada debe ir hacia las BBDD, no podemos pensar que es algo que está fuera de nuestra profesión”*, abogando también por el trabajo cooperativo al igual que Irune. Una muestra clara es el proyecto colaborativo explicado por la ponente, la **comunidad VALORA**. Es un proyecto promovido en su inicio por la **Asociación de Archiveros de Andalucía**, donde en la actualidad hay profesionales de toda España y en la que está muy presente también la valoración de la información contenida en bases de datos.

Después del descanso de la comida **Jordi Serra** retomó el testigo, esta vez para explicarnos su proyecto ya implantado en el **Consorcio de Educación de Barcelona**, donde trabaja como *Jefe de Sistemas de Información*. Según sus palabras: *“ Es un proyecto de gestión documental, porque en el fondo es un proyecto de transformación digital. No es un proyecto de archivo”*. Tal y como nos explicó a la mañana, se basa en un **modelo de gestión orientado hacia los datos**. *“Orientar una organización a datos no es solo tecnología, es transformar la forma de funcionar esa organización”*. Nos explicó que el primer elemento básico para que funcione cualquier política de gestión de datos o de gestión documental es el **soporte gerencial**. A partir de ahí Jordi nos guio a través de la transformación que se había producido en su institución. Una transformación que necesariamente deber ser rápida y aprovechando todo los recursos a su alcance. Aprendimos que la gestión de la información ha de entenderse como un servicio y que ese servicio solamente es efectivo cuando la información se gobierna de forma integrada. Entre tanta complejidad el punto positivo se vio cuando Jordi mostró gráficamente y con números contrastables los resultados obtenidos en el Consorcio. Jordi nos animaba en el camino: *“La organización informada tiende de manera natural a la mejora”*. Pero la organización solo está

informada cuando tiene datos que puede interpretar y no documentos que no puede procesar: *“Datos...a ser posible ciertos, consistentes, no redundantes y obtenidos de forma sostenible”*.

El modelo que implantaron en el consorcio se cimienta en **4 fases** *“que no se desarrollaron de forma secuencial, se desarrollan de forma continua. Cada día estamos identificando, cada día estamos transformando, estamos integrando y estamos sirviendo”*. La primera fase es la creación de un modelo de datos (modelo ontológico) con una visión holística de la organización (**IDENTIFICAR**). La segunda y tercera fase (**TRANSFORMAR** e **INTEGRAR**) consisten en la despatrimonialización de los datos. Hay que realizar una transformación rápida a través de la datificación de todos los procedimientos de la organización y desarrollar un **banco de datos** de la institución basado en el modelo ontológico de la primera fase. El objetivo es integrar todos los datos de la organización en un banco de datos alimentado de la manera más automatizada posible, *“la automatización es clave”* según Jordi. La última fase se refiere al proceso de repatrimonialización de los datos a través de los servicios ofrecidos a la propia institución (**SERVIR**). Jordi puso ejemplos de servicios construidos en el Consorcio, productos de información rápidos, personalizados e interactivos, tales como vistas estadísticas, cuadros de mando, servicios realizados con analítica predictiva, aplicaciones móviles, fichas de las entidades del modelo de datos con métricas calculadas, visualizaciones de movilidad a través de mapas, etc. El final de la última ponencia fue un viaje a un futuro no lejano, allí donde todos nos gustaría que estuviera nuestra organización: una organización informada y que tomara las decisiones a través de los datos obtenidos *“y no actuar de acuerdo con mi patrón, si no a la realidad que estoy viendo”*. Según el ponente todo ello no hubiera sido posible sin un equipo de trabajo altamente formado e integrado.

A lo largo de toda la jornada, y también antes del inicio de la mesa redonda, los asistentes presenciales y online formularon **diversas preguntas** a los ponentes. La interacción siempre es sinónimo de debate, reflexión e intercambio de ideas. A través de esas preguntas abordamos temas tan variados como los modelos ontológicos, la autenticidad de los datos, la propiedad vs responsabilidad de los datos, la protección de datos y la extracción de datos mediante la tecnología OCR, entre muchas otras preguntas. En la mesa redonda se abordó el futuro y presente de la profesión, donde hubo opiniones diferentes con un mensaje eminentemente optimista en cuanto a la preparación de los profesionales de la gestión documental y dispar en lo que se refiere al grado de madurez de las organizaciones.

Los ponentes terminaron con una **última reflexión** acerca de la jornada.

**Pilar** aludió en su conclusión final al tema de la preservación de los datos: *“Empecemos por algo, aunque sea algo sencillo. Y si nos equivocamos, nos hemos equivocado. Pero no tomar decisiones por no equivocarse, ya estáis tomando una decisión y es equivocada. Ya iremos mejorando el sistema poco a poco”*.

**Jordi** cogió el testigo de Pilar reforzando su conclusión: *“Como se aprende es en la práctica. Todas las metodologías consolidadas se han desarrollado a partir de práctica y de prueba-error. Tampoco hay que dar por sentado que lo que se desarrolla a nivel académico y de investigación es válido”*.



**Beatriz** resaltó la importancia de los entornos colaborativos y cooperativos *“No tenemos que estar inventando la rueda todo el día. Si otros lo han hecho, lo han hecho bien y lo han testado, podemos aprovecharlo”*.

**Irene** insistió en que la solución parte de grupos de trabajo colaborativos en las organizaciones y en que se debe intentar evitar que ocurra la afirmación de la siguiente cita anónima que afirma *“que la administración electrónica iba a ser el fin de la historia”*.

Los datos, opiniones y muestras de agradecimiento recogidas después de la jornada mostraron que las **expectativas** fueron cubiertas con creces: 80 personas acudieron presencialmente, 372 conexiones únicas el día de la emisión online y 185 nuevas visualizaciones en la semana posterior a la jornada.



La gestión documental en la era de los datos  
Valoración y preservación digital de datasets