

En Profundidad

La permanente transformación cultural, además de digital, para gestionar y compartir datos

Autor: Cap. D. Manuel Ángel de Pedro Cibanal (CIETO EA - TEL), Ministerio de Defensa, Secretaría de Estado de Defensa, Dirección General de Asuntos Económicos, Subdirección General de Contratación, Grupo de Estimación de Costes.

Palabras clave: gestión de datos, arquitectura de referencia única para la gestión de la información y del conocimiento del Ministerio de Defensa, AR GIC, transformación digital, *Nato Architecture Framework* versión 4, NAFv4.

Líneas I+D+i ETID relacionadas: Área 11.

Introducción

Podemos afirmar que todas las organizaciones son cada vez más conscientes de la importancia del análisis de sus datos. De hecho, en el actual contexto de incertidumbre y cambio constante, es crucial tanto para el Ministerio de Defensa como para la Alianza Atlántica tomar decisiones basadas en datos para la consecución de sus objetivos de una manera ágil y eficiente, no solo en el entorno operativo, también en sus aspectos de gestión interna. Este artículo trata de plantear una reflexión sobre modos y maneras al respecto.

Sin embargo, dar el salto para que personas y algoritmos de inteligencia artificial (IA) utilicen datos de calidad que asistan en la toma de decisiones, operativas o funcionales, requiere reevaluar el modelo de organización desde sus cimientos para ser, de verdad, Data Centric. Este término significa poner los datos en el centro de todas las actividades, otorgarles el protagonismo que merecen y abordar un cambio cultural en su gestión para poder compartirlos de manera automática con personas y «máquinas».

Para conseguirlo es imprescindible contar con capacidades analíticas

que capturen datos de calidad. Estos datos son el resultado de nuevas actividades de gestión desarrolladas en entornos que protegen los activos de información, sin que dicha protección afecte a su transmisión o tratamiento.

La gestión de datos

La gestión de datos (o *data management*) es la solución a las nuevas necesidades y permite poner a disposición conjuntos de datos (*datasets*) precisos, auténticos y fiables en cualquier momento de su ciclo de vida. Además, sus procesos facilitan la flexibilidad necesaria para que los usuarios autorizados, ya sean humanos o algoritmos, puedan acceder a los datos en formato digital de manera segura y ágil, en el momento y lugar adecuados (concepto de superioridad en la información).

Según la «Estrategia de desarrollo, implementación y uso de la Inteligencia Artificial en el Ministerio de Defensa», hemos entrado en una era en la que los *datasets* para el entrenamiento, junto con los modelos algorítmicos de IA, se deben convertir en activos estratégicos totalmente gestionados para su reutilización y puesta en valor.

Cualquier científico de datos puede corroborar las siguientes expresiones: «No confío en el análisis de datos de baja calidad porque no arroja resultados fiables», o «El 80 % de mi tiempo lo dedico a limpiar datos para poder utilizarlos y solo un 20 % lo destino al entrenamiento y análisis de modelos». Por lo tanto, es más necesario que nunca establecer no solo infraestructuras técnicas seguras para la gestión de *datasets*, sino también procesos de gestión previos al entrenamiento de cualquier modelo algorítmico, con el fin de facilitar el trabajo de los científicos de datos y de los analistas tradicionales.



Figura 1. La gestión de datos: origen de datos de calidad para su explotación por analistas e inteligencia artificial. (Fuente: *Food-For-Thought Paper on A Data-Centric Reference Architecture for the Alliance*. NATO C3B Data Management Capability Team).

Para evitar los problemas mencionados anteriormente, existen marcos de referencia de buenas prácticas de gestión de datos aceptados internacionalmente y que debemos aplicar cuanto antes. El Ministerio de Defensa ya ha dado cuando menos un primer paso al elaborar un documento llamado *Arquitectura de Referencia Única para la Gestión de la Información y del Conocimiento (AR GIC v1.0)*. Este documento, en formato de Arquitectura OTAN NAFv4, reúne las mejores prácticas en gestión de datos (*Data Management Book of Knowledge - DMBok*) y en el uso de infraestructuras tecnológicas que faciliten el movimiento de grandes conjuntos de datos

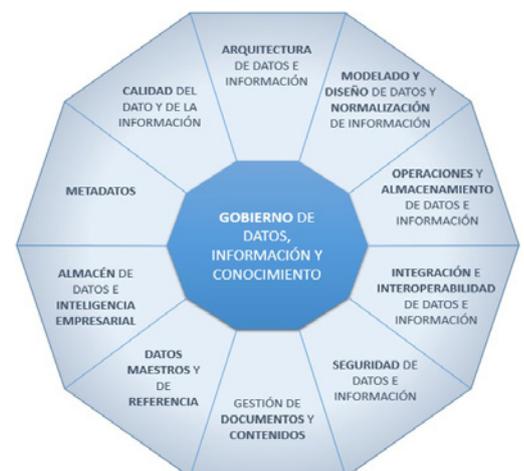


Figura 2. Capacidades GIC del MDEF. (Fuente: Vista C1 del Documento NAFv4 «Arquitectura de Referencia Única para la Gestión de la Información y del Conocimiento»).

En profundidad



Figura 3. Índice NAFv4 "Arquitectura de Referencia Única para la Gestión de la Información y del Conocimiento. (Fuente: Vista A2 del Documento NAFv4 «Arquitectura de Referencia Única para la Gestión de la Información y del Conocimiento»).

para su explotación (NIST *Big Data Interoperability Framework*).

Asimismo, esta primera versión del documento establece una hoja de ruta para alcanzar las once capacidades necesarias (con sus principios, procesos, productos de información y requisitos del *software* que apoyarán a dichos procesos) para lograr la adecuada gestión de los datos, la información y el conocimiento del Ministerio de Defensa.

Estas once áreas guían la gestión de datos remarcando la importancia de actividades como el mantenimiento de datos maestros y de referencia, la medición de indicadores de calidad de los datos, el uso correcto de datos estructurados o no estructurados, la aplicación de reglas de integración de datos, el uso de plantillas estandarizadas para la generación de documentos y/o analíticas (información), el uso de etiquetas de seguridad para la distribución segura y cifrada, la aplicación de protocolos de manipulación según qué tipo de dato, el meta datado semántico (no solo estructural) con esquemas de metadatos documentales y de diferentes comunidades de interés, el mantenimiento del linaje de datos que se van generando, la aplicación de procedimientos de acceso o la gobernanza de las taxonomías, términos y ontologías de las familias funcionales establecidas por la Estrategia de la Información del MDEF.

Características del documento y aspectos a mejorar AR GIC:

Las Arquitecturas NAFv4 (como la AR GIC), tienen un índice en formato

de tabla para facilitar su lectura. Dentro de cada área se usa un meta modelo que relacionan todas las entidades, garantizando la representación de conocimiento de la materia desde diferentes puntos de vista en un lenguaje común. No es el único documento de este tipo, ejemplos de esta nueva manera de redactar son las arquitecturas promocionadas por la Política CIS/TIC del MDEF, las Arquitecturas Operativas redactadas con la metodología NATO *Mission Thread* y la documentación de las Redes de Misión FMN (*Federated Mission Networking*) de la OTAN.

Este formato, en apariencia de fácil comprensión, no lo es para muchos roles clave del Ministerio. Los motivos: la inercia de documentar de manera tradicional, la reticencia a la aplicación de nuevos estándares y la reciente novedad del uso del Marco de Arquitecturas

OTAN (NAFv4) para redactar documentos interoperables dentro de la Alianza.

También se observa, en esta primera versión de la AR GIC, que el documento no se vincula del todo con las buenas prácticas de Gestión de la Información de la OTAN y sus normas homólogas de Gestión de Documentos Electrónicos de la AGE y del MDEF, junto con sus respectivos esquemas de metadatos publicados (ver bibliografía).

Este documento (no solo en su contenido, sino en su formato), comienza a hablar de conceptos con los que muchos actores internos del Departamento están todavía en fase alineación. Ya no se dejará a la interpretación lo que hay que hacer, y se deberá completar cada área del índice interrelacionándola con la normativa ya elaborada para generar una única fuente de verdad (del anglicismo *Single Source of Truth* -SSOT-). Para lograrlo, a partir de ahora existirán más roles que los tradicionales elaboradores, propietarios delegados y custodios (normalmente de documentos). Los nuevos roles como son el Oficial Jefe de la Información (CIO), el Oficial Jefe del Dato (CDO), las Autoridades Coordinadoras de Datos por Familia Funcional, los Arquitectos de Datos, los Custodios de Datos, los Curadores de Datos, etc. tendrán nuevas tareas y los antiguos roles tendrán que adaptar ciertos procesos de trabajo. Sus nuevas actividades diferirán significativamente de las tareas que venían llevando a cabo hasta ahora en relación con la gestión de datos (y de documentos electrónicos) y sus jefes orgánicos deberán de ser conscientes de que estos procesos serán la prioridad absoluta en muchos casos.

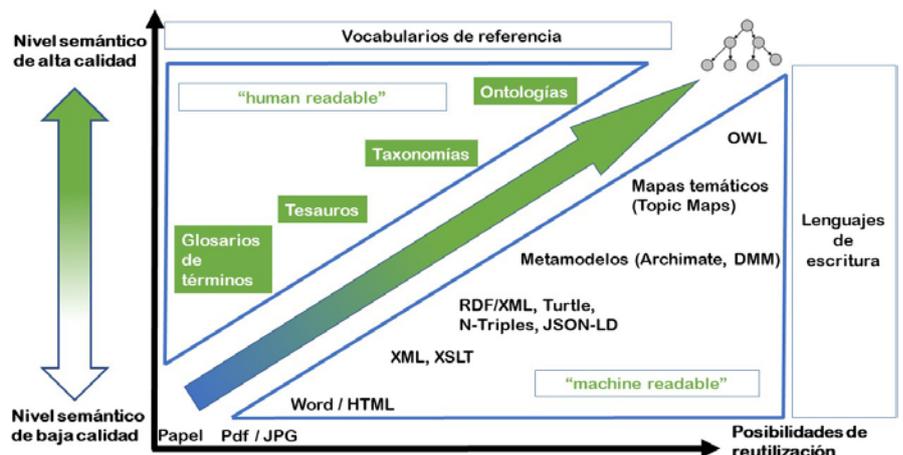


Figura 4. Evolución de la madurez semántica de los datos, la información y el conocimiento. (Fuente: propia).

Hay que añadir que la AR GIC es un documento estructurado, con vocabularios de referencia para ser leído por personas, pero la base de algo a lo que aún no estamos acostumbrados: *un documento machine readable*, formado por *Building Blocks (BB)*, análogos a los de las Taxonomías OTAN C3 y cuyos conceptos parten del estándar TOGAF de Arquitectura Empresarial (modelo que representa el negocio para determinar y gestionar la tecnología que lo apoya en una organización). Estos BB representan taxonomías de términos entrelazados (*Linked Data*) y meta datados semánticamente (*Web Semántica*), y su objetivo es la creación de un mismo lenguaje para que todas las partes interesadas puedan entenderse (única fuente de verdad para una comunicación que evite interpretaciones erróneas).

Quizá esta primera versión de la AR GIC debería haber sido escrita mejor en un lenguaje estandarizado de modelos empresariales como ArchiMate® (uno de los lenguajes de escritura que sugiere la documentación OTAN NAFv4), con las relaciones definidas entre sus entidades para facilitar su reutilización por otras personas y por herramientas de gestión de datos y algoritmos de IA. Este nuevo «documento electrónico», junto con otros, se integrará en el Modelo Semántico Digital del Ministerio de Defensa, formado por diferentes ontologías, según una de las actuaciones del Plan de Acción para la Transformación Digital del MDEF. Entender los anteriores puntos tiene sus dificultades sin la adecuada

«evangelización» y pudiera ser fuente de incompreensión e indiferencia hacia la AR GIC. Sin ser perfecto, este documento está adelantado a su tiempo y requerirá un periodo de aprendizaje para que las personas de la organización avancen de forma lo más fluida posible en su aplicación.

De manera adicional, los estándares anteriormente nombrados requieren ser usados por todo el Departamento (como ya ha empezado a hacer la OTAN), para un mejor entendimiento y facilitar la interoperabilidad. Pero como toda novedad, el esfuerzo adicional para aprender nuevas formas de trabajar/documentar hace necesaria

una adecuada gestión del cambio para el fomento de su aplicación y la comunicación de las ventajas que surgen de esta nueva manera de hacer las cosas.

Por lo tanto, y por lo anteriormente expuesto, será lógico que el Grupo de Trabajo colaborativo de Gestión de la Información y del Conocimiento (GIC) y de Transformación Digital inter-ámbitos, cuyas bases han sido establecidas por la Instrucción de Coordinación para la GIC y el Plan de Acción para la Transformación Digital, mejore en futuras versiones los aspectos antes mencionados para generar datos en forma de *Building Blocks* accesibles y compartidos.

La AR GIC v1.0 tiene la ventaja de ser una «combinación de conocimientos». Ahora tendrá que seguir evolucionando para facilitar su comprensión en futuras versiones y madurar semánticamente.

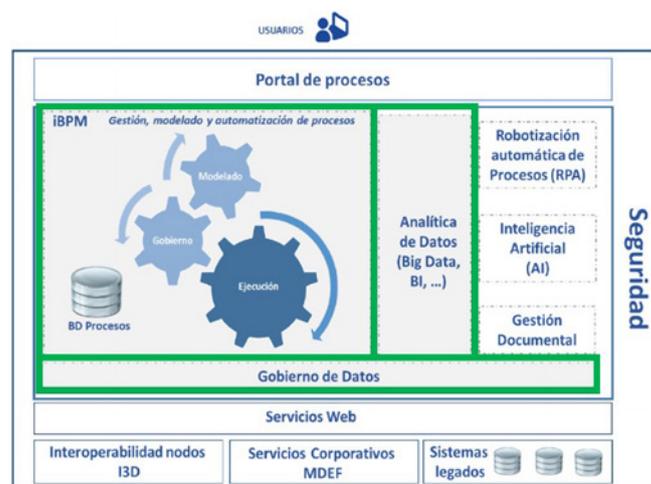


Figura 5. Primeros lotes licitados de la plataforma ARGO en verde. (Fuente: Grupo de trabajo inter-ámbitos del MDEF de Trasmformación Digital y GIC).

De manera adicional, y de ahí otro de los puntos importantes del documento, gracias a la hoja de ruta definida en la AR GIC v1.0, se están juntando las primeras piezas que ayudarán a gobernar los datos en la Plataforma de Armonización para la Gestión de la Organización (ARGO), donde el Departamento ha puesto grandes esperanzas de uso. Estas piezas son herramientas que forman parte del ya adquirido *IBM Cloud Pack 4 Data*, y que habilitarán a las aplicaciones (nuevas y legadas) del MDEF, a las nuevas herramientas analíticas de IA y a las personas un acceso seguro a los datos, siempre garantizando la

accesibilidad a la información, de una manera gobernada e interoperable.

Necesario un trabajo previo

Las empresas han analizado la necesidad del Departamento para su «modernización digital» y se están presentando a los diferentes lotes de ARGO. Gracias a ello, se están adquiriendo las herramientas más avanzadas en apoyo a los procesos de gestión de datos. De hecho, para ellas resulta muy sencillo vender las ventajas tecnológicas, el software avanzado y los beneficios de los tratamientos automatizados de datos que no requieren autorizaciones continuas (característica típica de organizaciones muy jerarquizadas).

Sin embargo, el Ministerio de Defensa no es ni una empresa, ni civil, ni al uso, y sus datos aún no están preparados para ser compartidos con otras personas o con agentes inteligentes de IA, sobre todo debido a la gran compartimentación orgánica que existe. Además, en estas primeras fases, y principalmente por este motivo, el uso de las nuevas herramientas tendrá un valor limitado, aunque ayudarán a iniciar, operar y completar un cambio de paradigma.

Bajo la ilusión de soluciones «todo en uno», las nuevas herramientas no accederán a ninguna base de datos si los creadores no quieren o no habilitan mecanismos, para que todo potencial consumidor acceda a «sus aplicaciones legadas o silos de información».

En otras palabras, por parte de la institución y de sus miembros sería clave gestionar los datos de una manera más adecuada y perfeccionada desde el inicio (aplicando las once áreas clave citadas de la AR GIC) con el fin de aprovecharlos y compartirlos de manera más eficiente. Pero de todo se aprende y se abre un nuevo mundo de oportunidades en el que es importante destacar que hay que tener en cuenta las necesidades de múltiples comunidades de interés generadoras de datos. De ahí la importancia del cambio cultural y de establecer mecanismos de gobernanza, para compartir y adoptar un nuevo modelo de organización, que aproveche al máximo todos sus datos y las nuevas herramientas que vayan llegando.

En profundidad

Beneficios

Sin entrar en la necesidad inevitable de alcanzar la Superioridad en la Información, para conseguir la Superioridad en el Enfrentamiento de las FAS (en el entorno multi dominio y en el *Combat-Cloud* a través de la interconexión de los datos de todo tipo originados por sistemas de armas), está demostrado que los nuevos tratamientos analíticos disponibles desde cualquier parte de la organización pueden agregar aún más valor. La fusión y análisis de diferentes fuentes de datos heterogéneas, que antes estaban separadas tanto tecnológica como organizativamente, van a servir para mejorar aspectos de funcionamiento de todo el Departamento.

La visión de sistemas aislados y usuarios exclusivos, pertenece al pasado. En la actualidad, todos los sistemas deben coexistir y compartir, aprovechando las nuevas oportunidades que ofrecen los flujos de datos a través Servicios CIS/TIC (como los llaman las FAS siguiendo las taxonomías de *Building Blocks/Términos C3* de la OTAN).

Por ello, hay que incidir en que los procesos de integración e interoperabilidad de los nuevos repositorios compartidos se conseguirán a través de metadatos semánticos, no solo estructurales, como resultado de la aplicación de la gestión de datos. Gracias a estos metadatos se facilitará exponencialmente la Ciencia de Datos, habilitando a los repositorios como fuentes de datos para entrenar algoritmos de calidad, sin sesgos, auditables y con reglas comunes, y la integración de nuevas aplicaciones orientadas al usuario basadas en la web semántica. De esta manera, los Servicios CIS/TIC accederán fácilmente a diferentes tipos de bases de datos (nuevas o legadas), a repositorios de información (documentos electrónicos) y a ficheros *machine readable* llenos de conocimiento (ontologías).

La nueva analítica de la IA solo podrá avanzar si los datos

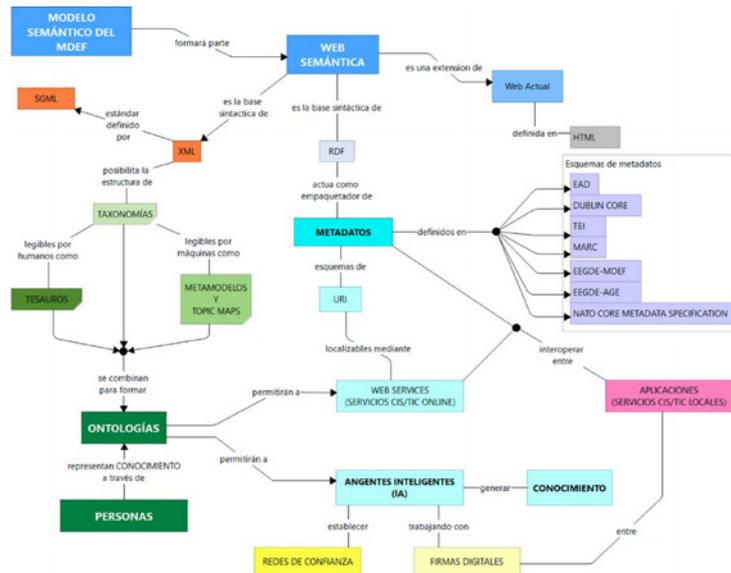


Figura 6. Esquema de la Web Semántica (Web 3.0). (Fuente: propia GIC).

se estructuran como activos permanentes y compartidos, y sería interesante que el Departamento siguiera, el ejemplo de la Alianza para ser *Data Centric a través de* los Objetivos VAULTIS: datos visibles (*Visible*), accesibles (*Accesible*), comprensibles (*Understandable*), vinculados (*Linked*), fiables (*Trustworthy*), interoperables (*Interoperable*) y protegidos (*Secure*). Esto supone un desafío cultural para los siete ámbitos del Departamento (SUBDEF, SEDEF, SEGENPOL, EMAD, ET, AR y EA) y para toda la orgánica que hay detrás. Los datos ya no son «tus datos», son datos del Ministerio de Defensa y son «los datos que se van a compartir» para satisfacer las necesidades que surjan en cualquier ámbito, orgánica o comunidad de interés (*Community of Interest, Col*) a través de una adecuada gobernanza.

De esto deriva que existirán nuevos usuarios que podrán utilizar datos para obtener nuevas visiones temporales y

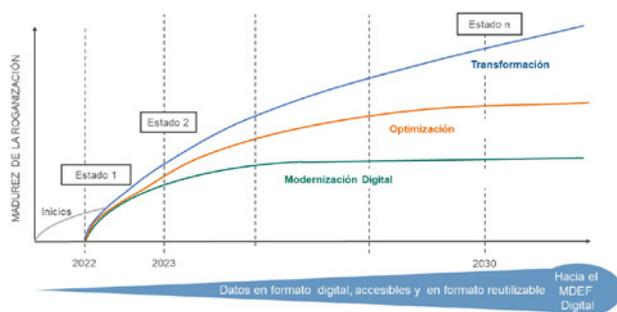


Figura 7. Hoja de ruta de la Transformación Cultural y Digital del MDEF. (Fuente: Presentación II Jornadas de IA del Ejército del Aire 2022, elaboración propia).

agregadas de los hechos a lo largo del tiempo con datos sin aparente relación. No solo eso, podrán compartirlos para contrastarlos y sugerir cambios para mejorar y «optimizar» toda la organización. Este es el objetivo último de una organización centrada en el conocimiento y las personas (Knowledge Centric Organization, KCO), dar a sus componentes la oportunidad de aportar y documentar su conocimiento para que no se pierda y para que contribuya a mejorar la organización de la que forma parte, para que se «transforme», para que sea más ágil y eficiente.

Cambio cultural

Todo lo anterior supone un cambio cultural importante, pasando de la necesidad de conocer (el famoso *need to know*), al deber de compartir de una manera segura (el nuevo *duty to share*). Todo ello dejando atrás la antigua práctica de restringir el acceso a los datos, fruto del pensamiento desactualizado de que «la información es poder y no la comparto».

Se considera necesario aplicar las mejores prácticas de gestión de datos, incorporando desde ya esas nuevas actividades que antes pudieran no considerarse necesarias o no eran realizadas de una manera estandarizada. El objetivo que se propone es avanzar más aún en convertir los datos en un activo estratégico y compartirlos para su uso de manera segura a través de la Infraestructura Integral de Información para la Defensa (I3D).

Llevará tiempo, pero forma parte del proceso de transformación permanente del Ministerio de Defensa (o de cualquier entidad en nuestros días) asimilar y aplicar plenamente las nuevas actividades de gestión de datos. Algunos objetivos que este autor ve positivos en la línea del artículo serían, por ejemplo, que el Departamento dispusiera de alguna forma de espacios de datos abiertos para compartir y reutilizar información de gestión entre diferentes ámbitos, universidad e

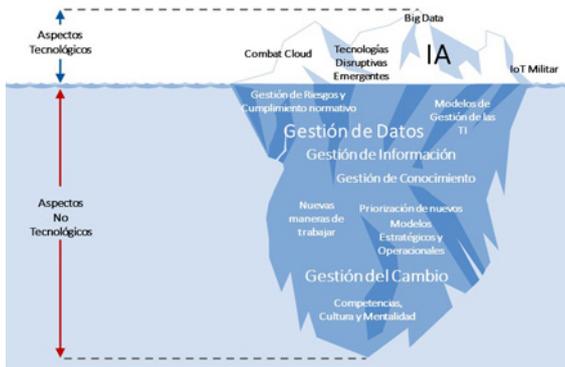


Figura 8: Aspectos tecnológicos y no tecnológicos de la Transformación Cultural y Digital del MDEF. (Fuente: Presentación II Jornadas de IA del Ejército del Aire 2022, elaboración propia).

industria para permitir descubrir cómo mejorar procesos internos, respetando sus peculiaridades, también compartidas a través de los modelos, meta datos y ontologías generadas por cada comunidad de interés; y de manera adicional abrir dentro la institución repositorios de datos operativos, para facilitar la interoperabilidad entre organismos de las FAS y aliados en ejercicios/zonas de operaciones.

Por último, seguir trabajando en la mejora de la agilidad, eficacia y eficiencia y en avanzar en la constante transformación del Ministerio de Defensa en una entidad que progresa basándose en el conocimiento aportado por las personas (ya sean actores internos o externos). Este es el principal recurso del Departamento y se debe aprovechar al máximo a través de las nuevas tecnologías digitales, no solo para lograr la debida adaptación al cambio, sino para estar siempre en vanguardia.

Bibliografía

- [1] ArchiMate Standard®. Disponible en: <https://pubs.opengroup.org/architecture/archimate3-doc/>
- [2] Arquitectura de Referencia única para la Gestión de la Información y el Conocimiento en el Ministerio de Defensa (AR GIC). Disponible en: <http://calderon.cud.uvigo.es/items/9b6d30af-090b-4d75-aa64-51e77474aa48/full>
- [3] *Data Management Book of Knowledge – DMBoK*. Disponible en: <https://www.dama.org/cpa-ges/body-of-knowledge>
- [4] *Department of Defense Data Strategy*. Disponible en: <https://media.defense.gov/2020/Oct/08/2002514180/-1/-1/0/DOD-DATA-STRATEGY.PDF>
- [5] Estudio y propuesta de uso del lenguaje «ArchiMate®» para generación de arquitecturas NAFv4 en el Ministerio de Defensa. Disponible en: <http://calderon.cud.uvigo.es/items/ec7a5337-375f-4907-8831-a4ced9dbad3d>
- [6] Instrucción 58/2016, de 28 de octubre, del secretario de Estado de Defensa, por la que se aprueba la Arquitectura Global de Sistemas y Tecnologías de Información y Comunicaciones del Ministerio. En: *Arquitectura global de Sistemas y Tecnologías de Información y Comunicaciones del Ministerio de Defensa (AG CIS/TIC)*. Disponible en: <https://publicaciones.defensa.gob.es/arquitectura-global-de-sistemas-y-tecnologias-de-informacion-y-comunicaciones-del-ministerio-de-defensa-ag-cis-tic.html>
- [7] *Linked Data*. Disponible en: <https://datos.gob.es/es/noticia/linked-data-como-modelo-de-datos>
- [8] *NATO Architecture Framework v4*. Disponible en: https://www.nato.int/cps/en/natohq/topics_157575.htm

- [9] *NATO C3 Taxonomy*. Disponible en: https://www.nato.int/cps/en/natohq/topics_157573.htm#:~:text=The%20C3%20Taxonomy%20is%20a%20model%20that%20represents,connecting%20NATO%27s%20Strategic%20Concept%20and%20Political%20Guidance%20
- [10] *NATO Core Metadata Specification. NATO Information Management Authority (NIMA)*. Disponible en: <https://nima.reach.nato.int/SitePages/Home.aspx>
- [11] *NATO Information Management Policy. NATO Information Management Authority (NIMA)*. Disponible en: <https://nima.reach.nato.int/SitePages/Home.aspx>
- [12] *NIST Big Data Interoperability Framework*. Disponible en: <https://www.nist.gov/>
- [13] Orden DEF/1196/2017, de 27 de noviembre, por la que se establece la Estrategia de la Información del Ministerio de Defensa. Disponible en: https://boe.es/diario_boe/txt.php?id=BOE-A-2017-14417
- [14] Orden DEF/2639/2015, de 3 de diciembre, por la que se establece la Política de los Sistemas y Tecnologías de la Información y las Comunicaciones del Ministerio de Defensa. Disponible en: <https://publicaciones.defensa.gob.es/politica-de-los-sistemas-y-tecnologias-de-la-informacion-y-las-comunicaciones-del-ministerio-de-defensa.html>
- [15] PAe. *Metadatos*. Esquema de metadatos para la gestión del documento electrónico. Disponible en: https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/Archivo_electronico/pae_Metadatos.html
- [16] Plan de Acción del Ministerio de Defensa para La Transformación Digital. Disponible en: <https://publicaciones.defensa.gob.es/plan-de-accion-del-ministerio-de-defensa-para-la-transformacion-digital-libro-pdf.html>

- [17] Política de Gestión de Documentos Electrónicos de la AGE. Disponible en: https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/Archivo_electronico/pae_Politica-de-gestion-de-documentos-electronicos.html
- [18] Política de Gestión de Documentos Electrónicos del MDEF. Disponible en: <https://publicaciones.defensa.gob.es/politica-de-gestion-de-documentos-electronicos-del-ministerio-de-defensa.html>
- [19] Real Decreto 521/2020, de 19 de mayo, por el que se establece la organización básica de las Fuerzas Armadas. Disponible en: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-5190
- [20] Resolución 11197/2023, de 29 de junio, de la Secretaría de Estado de Defensa, por la que se aprueba la Estrategia de desarrollo, implantación y uso de la Inteligencia Artificial en el Ministerio de Defensa. Disponible en: <https://publicaciones.defensa.gob.es/media/downloadable/files/links/2/0/20230706.pdf>
- [21] Resolución 420/17058/2018, de 7 de noviembre, de la Secretaría General Técnica, por la que se da publicidad al Esquema de Metadatos para la Gestión del Documento Electrónico en el ámbito del Ministerio de Defensa. Disponible en: <https://publicaciones.defensa.gob.es/esquema-de-metadatos-para-la-gestion-del-documento-electronico-ministerio-de-defensa-eemgde-mdef-version-1-0.html>
- [22] *The Open Group Architecture Framework*. Disponible en: <https://www.opengroup.org/togaf>