

Marco metodológico para la creación, implantación y mantenimiento de Sistemas de Gobierno de Datos

Ismael Caballero¹[0000-0002-5189-1427], Fernando Gualo² [0000-0002-7800-7902], y
Mario Piattini¹ [0000-0002-7212-8279]

¹ Alarcos Research Group, UCLM, Ciudad Real, Spain
Ismael.Caballero@uclm.es, Mario.Piattini@uclm.es

² DQTeam, Ciudad Real, Spain
fgualo@dqteam.es

Abstract. Para poder optimizar el rendimiento que las organizaciones esperan obtener de sus datos, necesitan tenerlos datos gobernados de acuerdo con sus estrategias organizacionales. Para ello, deben desarrollarse e implantarse Sistemas de Gobierno de Datos. El diseño, implantación y mantenimiento de estos sistemas requiere del despliegue y coordinación de diversos componentes. Pero en la realidad, hemos podido constatar que las organizaciones se encuentran con una serie de problemas que les impiden implantar sistemas de gobierno de datos que respondan a sus necesidades y expectativas. En este artículo se propone un marco metodológico – compuesto por un modelo de sistema de gobierno de datos y un proceso para su diseño, implantación y mantenimiento - alineado al estándar internacional ISO/IEC 38505 y basado en el Modelo Alarcos de Mejora de Datos (MAMDv3.0). El proceso (KCAE) consta de cuatro etapas: Conocimiento (para determinar lo que se sabe sobre los datos y lo que se requiere saber), Control (para determinar si el control ejercido sobre los datos a través de las políticas es suficiente de acuerdo con el objetivo requerido), Adaptación (para dirigir la organización hacia la transformación del control actual hacia el control requerido), y Explotación (para supervisar si las adaptaciones y el control son suficientes para lograr los objetivos empresariales deseados). Finalmente se describe cómo se usó el marco para el caso concreto de un servicio en una universidad española.

Keywords: Gobierno de datos, Sistema de Gobierno de Datos, KCAE, ISO/IEC 38505, Modelo Alarcos de Mejora de Datos (MAMD)

1 Introducción

Los datos son un elemento fundamental para el desarrollo de las organizaciones, ya que, son el ingrediente fundamental de sus procesos de negocio, y pueden ser usados para mejorar la eficiencia en el uso de los recursos y para lograr un aumento de la productividad [5, 19, 25, 27]. Esto convierte a los datos en activos organizacionales, y como otros activos, deben ser gobernados convenientemente. Se puede decir, por tanto, que el gobierno de los datos se ha convertido en un habilitador de la economía de los datos [9], así como también en un pilar básico para la mejora de la transparencia y

eficiencia de las administraciones públicas [21, 22]. Para potenciar la economía del dato, la UE ha desarrollado la Estrategia Europea de Datos [11], con la que espera proyectar para el año 2030 un Espacio Único Europeo de Datos que fomente un ecosistema con nuevos productos y servicios basados en el uso, compartición y explotación de los datos. Sin embargo, para alcanzar este objetivo, la Estrategia Europea de Datos reclama la necesidad de mejorar las estructuras o sistemas de gobierno de datos de las organizaciones y satisfacer aspectos regulatorios tanto genéricos (como las leyes europeas GDPR o *Data Governance Act*, o las españolas ENS o ENI) como aquellos relacionados con sectoriales específicos (como Solvencia II para el sector seguros, o Basilea para el sector financiero). Con lo que se plantea como una necesidad para las organizaciones diseñar adecuadamente y establecer sistemas de gobierno de datos.

En cualquier caso, el establecimiento de sistemas de gobierno de datos presenta numerosos problemas y desafíos que hacen que su implementación sea una tarea compleja. Participando en experiencias reales en diversos tipos de organizaciones públicas y privadas de distinto tamaño, hemos constatado la presencia habitual de una serie de problemas: (1) falta un entendimiento adecuado de los conceptos de gobierno de datos, y de las implicaciones de tener los datos gobernados tanto al nivel de negocio como a nivel de gestión de datos; (2) inexistencia de estrategia de datos alineada a la estrategia de negocio que permita integrar los resultados obtenidos dentro de su sistema productivo; (3) dificultad de comprensión y cuantificación del valor de los datos; (4) apuesta exclusivamente por herramientas específicas que dan soporte tecnológico a algunos de los artefactos típicos (p.ej. glosario de términos, o diccionario de datos); (5) no involucración de la gente de negocio ni de dirección; (6) no hay una separación clara entre las responsabilidades de gobierno de datos y las de gestión tecnológica de datos; (7) insuficientes recursos y (8) fallos en los modelos y estándares existentes.

Por tanto, creemos que las organizaciones se podrían beneficiar de la existencia de un marco metodológico para la creación, implantación y mantenimiento de sistema gobierno de datos que intentará paliar los problemas anteriores. Este marco incluye dos líneas de acción: (1) la identificación de los elementos más representativos para desplegar un sistema de gobierno de datos y (2) la definición de un proceso para el diseño, implantación y mantenimiento del sistema de gobierno de datos adaptado a la realidad de las organizaciones donde se implantará. El proceso consta de cuatro fases: Conocimiento, Control, Adaptación y Explotación. El marco desarrollado se alinearán con los principales estándares internacionales abiertos relacionados con el gobierno de los datos (ISO/IEC 38505-1, ISO/IEC 38505-2, ISO/IEC 38505-3) y con el marco de referencia Modelo Alarcos de Mejora de Datos [8].

Este artículo está estructurado como sigue: en la sección 2, se muestran los aspectos más importantes del estado del arte; en la sección 3 se presenta el marco metodológico; en la sección 4 se describe cómo se usó el marco en el caso de un servicio de una universidad española, y finalmente, en la sección 5, presenta una serie de conclusiones y trabajos futuros.

2 Estado del arte

Salvo COBIT 2019 [13] y de alguna manera [27], la inmensa mayoría de los marcos académicos y profesionales (p.ej.[4, 5, 7, 10, 13, 15, 19, 21, 26]) no identifican explícitamente el concepto de “**sistema de gobierno de datos**”, sino que se establece bajo un paraguas más genérico de “gobierno de datos”. En cualquier caso, la idea es la misma. De entre los marcos incluidos en la literatura, y dada su importancia nos gustaría resaltar especialmente tres por su relevancia para este trabajo: (1) COBIT 2019 [13], que identifica para el sistema de gobierno de tecnologías y de información los siguientes elementos: procesos; estructuras organizacionales; principios, políticas y marcos de referencia; información; cultura, ética y comportamiento; personas, habilidades y competencias; servicios, infraestructuras y aplicaciones; además, COBIT propone una guía de diseño de sistemas de tecnologías de información [12] y una guía de implantación de los sistemas diseñados [14]. Otro de los marcos importantes es (2) ISO 38505-1[15] e ISO 38505-2 [16] que muestra los aspectos claves del gobierno de datos (**valor de los datos, riesgo, y restricciones**) e introducen seis principios (responsabilidad, estrategia, adquisición, rendimiento, cumplimiento y comportamiento humano) junto con unas guías de desarrollo teniendo en cuenta una serie de factores que determinan su implementación; identifica una serie de procesos (**evaluar, dirigir, monitorizar**) como áreas propias de actuación del gobierno de datos, distinguiéndolos y estableciendo las relaciones correspondientes con las operaciones propias de la gestión de datos. Sin embargo, no describe actividades específicas para la creación de sistemas de gobierno de datos. Finalmente, incluimos (3) el Modelo Alarcos de Mejora de Datos (MAMD) [8] por sus dos componentes (3a) un modelo de referencia de procesos (véase **Fig. 1**) que agrupa en tres categorías los procesos de gestión de datos (DM), gestión de calidad de datos (DQM) y gobierno de datos (DG); y (3b) un modelo de evaluación de procesos, que incluye un modelo de madurez organizacional por capacidades de los procesos que permitirá a las organizaciones priorizar los objetivos específicos de un sistemas de gobierno en función de sus fortalezas, sus debilidades y sus necesidades.



Fig. 1. Modelo de Referencia de Procesos de MAMD v3.0 (disponible para descarga gratuita de <https://mamd.dqteam.es>)

Un análisis de los marcos de gobierno de datos existentes permite llegar a la conclusión de que adolecen de varios problemas: (1) muchos de ellos son privativos (p.ej. DMM), normalmente con costes de licenciamiento muy altos; (2) están pensados para grandes organizaciones en sectores específicos -como banca-(p.ej. DCAM), y suelen requerir el despliegue de una considerable infraestructura organizativa no siempre al alcance de empresas más pequeñas; (3) son excesivamente genéricos, normalmente muy complejos, y con grandes diferencias entre ellos desde el punto de vista conceptual, lo que dificulta su implementación (p.ej. DAMA); (4) no existe un consenso sobre las áreas de actuación que entran bajo el paraguas del gobierno de datos; (5) salvo muy pocas propuestas, en general, estos marcos no abordan la idea de sistema de gobierno de datos, ni de los componentes que deberían tener, ni proporcionan aspectos metodológicos riguroso y sistemáticos para su creación. Sin embargo, es interesante mencionar que MAMD fue creado para evitar estos problemas, y, por tanto, la propuesta presentada en este artículo se centra en este modelo.

Los problemas anteriores se acaban traduciendo en iniciativas ad hoc para sistemas de gobierno de datos, haciéndolos difícilmente mantenibles y poco interoperables con los de otras organizaciones.

3 Marco metodológico para Sistemas de Gobierno de Datos

Como se indicó en la introducción, el marco metodológico consta de dos componentes fundamentales: (1) un modelo de sistema de gobierno de datos donde se identifican los elementos y los artefactos relacionados, y (2) un proceso de diseño, implantación y mantenimiento del sistema de gobierno de datos que tiene en cuenta determinados factores de diseño, así como el grado de madurez que es deseable alcanzar. Así, cuanto más maduro sea el sistema de gobierno de datos, más datos abarcará, más procesos típicos de gobierno de datos, gestión de datos y gestión de calidad de datos cubrirá y mayor será su nivel de capacidad. En esta investigación, se tomará MAMD como referencia tanto para los procesos citados como para describir los niveles de capacidad.

3.1 Modelo de Sistema de Gobierno de datos

COBIT 2019 identifica para un sistema de gobierno de Tecnologías de Información una serie de tareas específicas en cuatro dominios de actuación. Se pueden particularizar las tareas específicas de estos dominios en función del tipo de valor organizacional de los datos [20] que el sistema de gobierno de datos pretenda optimizar:

- **Alineamiento estratégico**, en los que se pretende conseguir que los datos estén alineados a la estrategia organizacional [1]. Con esto, y dependiendo del tipo de estrategia de datos planteada (p.ej. ofensiva vs defensiva [6]), se puede pretender maximizar cualquiera de los tipos de valores intrínsecos, de negocio o de mercado.
- **Realización de beneficios**, cuyo objetivo es asegurar que los datos puedan ser entregados a los consumidores en condiciones óptimas para su uso y explotación [10]. Con esto, lo que se pretende es maximizar el valor de negocio de los datos, o el valor

de rendimiento de los datos o el valor de mercado de los datos, dependiendo del tipo de estrategia de datos establecidas.

- **Optimización de riesgos**, donde el objetivo es asegurar que no se viola ninguna de las restricciones (p.ej. legales) pertinentes que pueden poner en peligro ni el uso de los datos ni la organización [19]. Por tanto, lo que se pretende con esta actuación es minimizar el valor perdido de los datos.
- **Optimización de recursos**, cuyo objetivo es conseguir que la gestión, uso, y explotación de los datos se realice de forma equilibrada y sostenible por parte de la organización usando la combinación adecuada de recursos [4, 10, 19]. En este caso, lo que se pretende es maximizar el valor económico de los datos.

También, tomando como base la estructura de sistema de gobierno de Tecnologías de Información propuesto en COBIT 2019 [13], en esta investigación se propone la siguiente estructura para un sistema de gobierno de datos (se incluyen los elementos, y se indican sus principales artefactos como productos de trabajo típicos):

1. **Procesos de DM/DQM/DG**, que se refiere a la particularización e implantación de los procesos de MAMD [8] (véase Fig.1). El número de procesos de MAMD elegidos y el nivel de capacidad a implantar deberá tener en cuenta la realidad de las organizaciones, considerando sus excepciones y las adaptaciones pertinentes (algo que se materializará a través de la idea de madurez de sistema de gobierno de datos). Cada proceso se describirá por medio de una plantilla que recogerá toda la información necesaria con respecto a la naturaleza del proceso y con respecto al nivel de capacidad que se desee o necesite alcanzar para dicho proceso. Esta descripción se apoyará en la del resto de los elementos del sistema de gobierno de datos particular de una organización. El principal artefacto de este componente es un cuerpo de conocimiento materializado típicamente como un repositorio de información (p.ej. una wiki) con la descripción de cada uno de los procesos seleccionados, y la relación entre ellos. Este repositorio estará disponible para todos los implicados de la organización, con objeto de facilitar una cultura de datos organizacional.
2. **Estructuras organizacionales**, en las que se incluye las cadenas de responsabilidades y rendición de cuentas [23], haciendo una separación entre gobierno de datos y gestión de datos. Los principales artefactos de este componente son por un lado (1) un catálogo de puestos de trabajo que incluya roles tales como jefe de datos (*Chief Data Officer*), Data Stewards, Propietario de datos, y (2) por otro, unas matrices RACI que establecen la relación de control, colaboración y comunicación entre los distintos roles. Roles y responsabilidades deberían estar presente en la descripción de los procesos DM/DQM/DG. Puede ocurrir que, en lanzamiento de las iniciativas de lanzamiento de la implantación de un sistema de gobierno de datos, se establezca un equipo inicial de soporte que luego acaben convirtiéndose en los que desempeñen los mencionados roles. Pero no necesariamente tiene que ocurrir esto.
3. **Principios, políticas y marcos de referencias**: que incluyen todos los principios (p.ej. los principios GAIP [19]), las directrices o políticas y los controles correspondientes asociados necesarios para modelar y gestionar el valor de los datos, el riesgo a asumir, y las restricciones a considerar, tal como se describe en ISO 38505-2 [16]. El control a ejercer sobre los activos de datos será eficiente cuando las políticas se

consensuen verticalmente entre todos los estamentos identificados en la estructura organizacional, pero aprobadas a alto nivel ejecutivo para que se pueda asignar la responsabilidad clara sobre los activos de datos. El artefacto más representativo de este componente es un repositorio de información (p.ej. una wiki interna) conteniendo el conjunto de políticas de datos asociadas a cada uno de los activos de datos, los respectivos controles que acrediten el cumplimiento de la política, y los correspondientes marcos de referencia de auditoría. Es importante tener en cuenta el tiempo de vigencia de la política, pues el valor de los datos es siempre relativo al tiempo de vigencia de las políticas. Para una gestión más eficiente de los activos de datos, suele ser recomendable crear una clasificación de etiquetas de los mismos, asociar políticas a cada etiqueta, y finalmente etiquetar los activos de datos en función de los criterios de la clasificación.

4. **Datos**, que incorpora una identificación de los activos de datos que deben ser gobernados junto con la descripción adecuada de los mismos. Esta descripción se realizará por medio de diferentes tipos de metadatos relevantes agrupándolos en diferentes repositorios de datos según su naturaleza [5, 7]: (1) metadatos de negocio, recogidos en un glosario de términos (p.ej. definición consensuada de los datos, posibles datos de referencia, y clasificación o categorías de los datos con respecto a determinados criterios como la sensibilidad), (2) los metadatos técnicos, recogidos en un catálogo de datos (p.ej. descripción de los repositorios de datos donde se alojan los datos incluidos en el glosario de términos); y (3) los metadatos tecnológicos recogidos en un diccionario de datos (p.ej. incluye detalles de implementación de los datos, tales como tipo y tamaño, restricciones, ...). Estos repositorios de metadatos serán convenientemente compartidos con los implicados, estableciendo criterios de acceso según la sensibilidad de los datos. Los artefactos que son principales productos de trabajo de este componente son los distintos repositorios de metadatos (glosario de negocio, catálogo de datos, y diccionario de datos).
5. **Cultura, ética y comportamiento**, que incluye la identificación de aspectos éticos (p.ej. aquellos relacionados con determinados algoritmos IA para evitar diferentes tipos de sesgos[17]), y aspectos relacionados con la cultura organizacional de datos también conocida como *data literacy* [18]. El principal componente de este elemento es un conjunto de directrices e iniciativas de sensibilización organizacional para diferentes aspectos de los datos (p.ej. garantizar la calidad o la seguridad de los datos de forma vertical, o consultar los repositorios de metadatos para cualquier acción).
6. **Personas, habilidades y competencias**, que incluiría detalles sobre los requisitos que deberían satisfacer las personas que desempeñen los roles implicados en la gestión y uso de datos, gestión de calidad de datos y gobierno de datos [23]. Los principales artefactos de este componente son (1) un listado de conocimientos relacionados con los diferentes procesos, (2) gestión de carreras profesionales, (3) planes de formación de competencias, (4) planes de adquisición de talento y (5) planes de contingencia ante la eventual rotación funcional de talento.
7. **Servicios, infraestructuras y aplicaciones**, que aborda todo lo relacionado con las TI necesarias para dar soporte a la gestión de los datos como base para los procesos de negocio [3, 10]. También se incluyen aquellas aplicaciones que dan soporte a operaciones específicas de gobierno de datos, como aquellas relacionadas con la

gestión de políticas, o las de soporte a los diferentes repositorios de metadatos identificados [19, 24]. El artefacto principal es un catálogo de servicios de datos, infraestructuras tecnológicas y aplicaciones para la adecuada gestión, gestión de calidad de datos y gobierno de los datos.

3.2 Proceso KCAE.

El objetivo del proceso KCAE (acrónimo de Conocimiento, Control, Adaptación y Explotación) es diseñar, implantar y mantener un sistema de gobierno de datos con el nivel de madurez adecuado y suficiente como para conseguir los objetivos de gobierno de datos de la organización. A través de experiencias de uso reales y haciendo una analogía con el Proceso Unificado de Desarrollo, hemos podido constatar que este proceso tiene estas tres características:

1. **Es iterativo e incremental**, pues es en sí mismo un programa de mejora continua de la forma en la que la organización aborda la forma en la que gestiona y usa los activos de datos. Esta mejora se produce en dos direcciones: (1) abarcando cada vez más procesos de gestión, gobierno y uso de los datos, y (2) estableciendo niveles de capacidad más exigentes para dichos procesos. En este sentido, el modelo de madurez proporcionado por MAMDv3.0 puede ayudar a entender cómo proyectar la mejora. La complejidad del sistema de gobierno de datos aumenta a un ritmo rápido a medida que se integran más activos de datos, se incluyen más procesos a niveles superiores de capacidad y se consigue involucrar a más funciones del negocio para usar los datos gobernados [2]. Para gestionar esta complejidad es imprescindible diseñar un plan director que recoja cómo la organización debería ir desarrollando progresivamente y de forma sostenible los elementos del sistema de gobierno de datos a través de la concatenación de programas de gobierno de datos, valorando convenientemente el riesgo del desarrollo. Por eso, muchas organizaciones están creando lo que se viene en llamar “oficina de datos” que tiene la responsabilidad de abordar de forma integrada los programas necesarios para el diseño, implantación y mantenimiento de su sistema de gobierno de datos.
2. **Está guiado por el valor de los datos**. La declaración del valor requerido de los activos de datos se realiza mediante el establecimiento de estrategias de datos [1]. Este valor deseado marcará el conjunto y el orden de las iteraciones e incrementos necesarios para construir el sistema de gobierno de datos. Estos incrementos tienen como objetivo que las organizaciones consigan que los activos de datos alcancen los niveles establecidos para los tipos de valores de datos [20], aunque habitualmente no sepan determinar ni proyectar de forma explícita dichos tipos de valores. Para monitorizar si se está consiguiendo alcanzar el valor de los datos deseados, es necesario establecer una serie de indicadores claves de procesos.
3. **Está centrado en la arquitectura empresarial**: la implantación del sistema de gobierno de datos es en sí mismo una transformación de la arquitectura empresarial de la organización a medida que se van ejecutando los programas de gobierno de datos incluidos en el plan director [10, 19]. Es importante notar que esta transformación tiene un doble objetivo: crear propiamente el sistema de gobierno de datos y dotar a la

organización de las capacidades necesarias para mantener o evolucionar de forma eficiente el sistema de gobierno de datos en el futuro. Para monitorizar el crecimiento se puede considerar el concepto de nivel de madurez del sistema de gobierno de datos. En este sentido, nuevamente, el modelo de madurez propuesto en MAMD puede ayudar a entender el grado de madurez al que el sistema de gobierno de datos debería llegar.

Teniendo en cuenta estas características, y antes de describir las etapas KCAE del proceso, es preciso establecer algunos aspectos básicos:

1. El diseño inicial del sistema de gobierno de datos dependerá de una serie de factores (tamaño de la organización, objetivos del negocio, ...). En este artículo no se abordarán dichos factores de diseño por restricciones de espacio.
2. El desarrollo de un sistema de gobierno de datos suficientemente maduro y evolucionado puede requerir varios ciclos KCAE a lo largo del tiempo, ciclos que se pueden proyectar en función de la estrategia de datos. Cambios en la estrategia de datos se pueden producir por voluntad propia de la organización (p.ej. necesidad de nuevas líneas de negocio) o bien motivada por cambios externos (p.ej. publicación de nueva regulación que ha de cumplirse).
3. Para la aplicación del primer ciclo KCAE debe formarse un equipo inicial que disponga del conocimiento adecuado, experiencia, credibilidad y autoridad suficiente como para lanzar el proyecto. Puede ser también necesario contar con la ayuda de una consultora especializada para dar los primeros pasos. Como consecuencia de esta primera iteración debería haberse establecido al menos una estructura organizacional suficiente que permita la aplicación de sucesivas aplicaciones KCAE. Es razonable pensar que miembros del equipo inicial acaben desempeñando los roles establecidos para el gobierno de datos en las iteraciones sucesivas [23].
4. Es recomendable evaluar el estado inicial (AS_IS) y el estado deseado del sistema de gobierno de datos (TO_BE) en términos de niveles de madurez proporcionado por algún modelo de referencia (en KCAE, el modelo de referencia será MAMD [8]). Teniendo en cuenta la estrategia de datos, se desarrollará un plan director para cerrar las brechas entre el estado final deseado y el estado inicial. Este plan director tiene que dar lugar a un conjunto de programas de gobierno de datos que, de forma iterativa e incremental, vaya desarrollando y/o manteniendo los diferentes elementos del sistema de gobierno de datos. En esta investigación, el hilo conductor del programa de gobierno es la implantación de los procesos de MAMD hasta el nivel de capacidad requerido. La consecuencia de la implantación de un proceso, además de la propia definición del proceso, es la creación o revisión de los artefactos o productos de trabajo de los diferentes componentes del sistema de gobierno de datos. Para ello, en las siguientes secciones se propondrán unas tareas básicas para cada etapa KCAE considerando lo mínimo que debería alcanzarse para cada nivel de capacidad, explicitando qué productos de trabajo se deberán generar. Por cada tarea se detallarán los procesos de MAMD en que son intensivos.
5. La ejecución de cada una de las etapas de KCAE requiere la colaboración de roles desempeñando responsabilidades de gobierno de datos para la toma de decisiones colegiadas y de roles desempeñando tareas de gestión de datos para proporcionar la información sobre la que se toman las decisiones.

3.3 Descripción de las etapas

En las etapas K y C (Conocimiento y Control) se recopila toda la información sobre el estado actual y deseado del conocimiento sobre los activos de datos, sobre la organización, así como sobre el nivel de control que se requiere establecer sobre los activos de datos para alcanzar los objetivos deseados. Estas etapas se corresponden con la tarea Evaluar de ISO 38505-1/COBIT 2019. En estas etapas, los roles de gobierno determinan las necesidades de conocimiento y control y los roles de gestión proporcionan la información necesaria para determinar si se están satisfaciendo dichas necesidades. Para lograr este objetivo, y en función del nivel de madurez que se desee alcanzar, se plantea la ejecución de las tareas mostradas en la segunda columna de la Tabla 1 y de la Tabla 2 para generar o modificar los productos de trabajo típicos en la tercera columna de dichas tablas. El acrónimo NM en la primera columna significa “Nivel de Madurez”

Tabla 1. Tareas y productos de trabajo típico de etapa K de Conocimiento

NM	Tareas	Productos Trabajo típicos
2	<ul style="list-style-type: none"> • K.01. Identificar y describir los procesos de negocio prioritarios (DM.01). • K.02. Identificar y describir las tecnologías de información que dan soporte a los procesos de negocio y a la explotación de los datos (DM.02) • K.03. Identificar y describir los activos de datos usados en los procesos de negocio a ser gobernados (DM.03) • K.04. Describir los niveles de seguridad alcanzados para los datos (DM.04) • K.05. Evaluar de los niveles de calidad de los activos de datos (DQM.01) • K.06. Listar y describir las políticas y estándares existentes para los datos (DG.01) 	<ul style="list-style-type: none"> • Mapa de procesos de negocio • Inventario de Tecnologías de información. • Glosario de Términos • Catálogo de datos • Diccionario de datos • Evaluaciones de calidad de datos • Descripción de los niveles de seguridad de los datos.
3	<ul style="list-style-type: none"> • K.07. Describir los modelos de datos (DM.07) • K.08. Describir las fuentes y destinos de los datos (DM.08) • K.09. Describir los procesos de integración de los datos (DM.09) • K.10. Identificar los posibles datos maestros (DM.10) • K.11. Describir del ciclo de vida de los activos de datos (DG.03) 	<ul style="list-style-type: none"> • Modelos conceptuales de datos • Mapas de fuentes y destinos de los datos • Inventario procesos ETLs • Mapas de roles y responsabilidades • Ciclo de vida de los datos
4	<ul style="list-style-type: none"> • K.12. Determinar el valor organizacional inicial de los activos de datos (DG.06) 	<ul style="list-style-type: none"> • Valor organizacional datos

En la etapa **A (Adaptación)**, tras un análisis DAFO realizado sobre los niveles de conocimiento y control actuales y deseados, se desarrolla un plan director en donde se proyectarán y desarrollarán los programas necesarios para obtener los incrementos necesarios. Esta etapa A se corresponde con la tarea **Dirigir** de ISO 38505-1/COBIT 2019. Los incrementos proyectados pueden estar orientados a evolucionar e integrar

uno o varios productos de trabajo de los componentes del sistema de gobierno de datos, y puede abarcar aspectos tan diferentes como diseñar y ejecutar un plan de formación para los trabajadores, crear y poblar inicialmente un glosario de términos, hacer una evaluación de calidad de datos para complementar el conocimiento que se tiene sobre un determinado activo de datos, o desarrollar una serie de políticas y controles para garantizar el cumplimiento de una normativa.

Tabla 2. Tareas y productos de trabajo típico de etapa C de Control

NM	Tareas	Productos Trabajo típicos
2	<ul style="list-style-type: none"> • C.01. Determinar si se tienen los mecanismos adecuados para la trazabilidad de los datos (DM.04) • C.02. Identificar la legislación genérica o específica del dominio de actuación de los datos, así como de los organismos para los que haya que reportar (DG.01) • C.03. Identificar los estándares sectoriales o tecnológicos cuya implementación beneficie a la organización (DG.01). • C.04. Listar todas las políticas de datos existentes (DG.01) • C.05. Comprobar que las políticas cubren legislación, estándares, normativas (DG.01) • C.06. Listar todas las etiquetas de clasificación de los datos (DG.01) • C.07. Determinar si las etiquetas de los datos agrupan bien las políticas de datos (DG.01). • C.08. Determinar qué activos de datos están cumpliendo las políticas de datos (DG.01) 	<ul style="list-style-type: none"> • Linaje de los datos • Listado de legislación a satisfacer • Listado de organizaciones a las que reportar • Listado de estándares sectoriales o tecnológicos a implementar. • Listado de políticas de datos • Glosario de términos enriquecidos con las etiquetas de datos.
3	<ul style="list-style-type: none"> • C.09. Comprobar el grado de cumplimiento de la estrategia de datos (DG.01) • C.10. Determinar si el ciclo de vida de los datos cubre las necesidades de la estrategia (DG.03) • C.11. Identificar los conocimientos y habilidades existentes y necesarios para cada una de las responsabilidades (DG.04) • C.12. Identificar las responsabilidades de cada uno de los roles (DG.05) • C.13. Determinar si las responsabilidades son suficientes para satisfacer estrategia de datos (DG.04) 	<ul style="list-style-type: none"> • Informe de cumplimiento de estrategia de datos. • Cadenas de responsabilidad y rendición de cuentas • Ciclo de vida de los datos • Competencias, habilidades y conocimientos para las responsabilidades
4	<ul style="list-style-type: none"> • C.14. Determinar si se han establecido los mecanismos necesarios para asegurar la calidad de los datos • C.15. Determinar si el valor de los datos de los datos es suficiente 	<ul style="list-style-type: none"> • Mecanismos de mejora de los datos • Valor de los datos

La ejecución de esta etapa requiere la colaboración de roles de gobierno de datos (para establecer, priorizar y validar los incrementos del sistema de gobierno de datos en función del valor que le puede proporcionar a la organización como concededores del apetito de riesgo, o la búsqueda de recursos para aprovisionar los recursos necesarios para la ejecución de los programas) y los roles de gestión de datos para ejecutar cada uno de los proyectos del programa de gobierno de datos derivados para conseguir cada

uno de los incrementos. Las tareas correspondientes para esta etapa se muestran en la Tabla 3.

Tabla 3. Tareas y productos de trabajo típico de etapa A de Adaptación

NM	Tareas	Productos Trabajo típicos
2	<ul style="list-style-type: none"> • A.01. Desarrollar los repositorios de metadatos, tanto a nivel tecnológico, como a nivel de contenidos (DM.01/DM.03) • A.02. Facilitar el acceso de usuarios y aplicaciones a repositorios de metadatos (DM.01/DM.02) • A.03. Desarrollar las políticas y controles de datos (DG.01) 	<ul style="list-style-type: none"> • Repositorios revisados de metadatos • Inventario revisado de tecnologías de información • Inventario revisado de políticas y controles de datos
3	<ul style="list-style-type: none"> • A.04. Revisar la estrategia de datos (DG.01) • A.05. Revisión del ciclo de vida de los datos (DG.03) • A.06. Establecer planes de formación continua para la adquisición de competencias, habilidades y conocimientos (DG.04) • A.07. Revisar la estructura organizacional (DG.05) • A.08. Revisión de cadenas de responsabilidad y rendición de cuentas (DG.05) • A.09. Desarrollar planes de comunicación entre roles y responsabilidades (DG.05) • A.10. Establecer planes de recompensas (DG.04/DG.05) • A.11. Desarrollar sistemas de resolución de conflictos (DG.05) 	<ul style="list-style-type: none"> • Estrategia de datos revisada • Ciclo de vida revisado de los activos de datos organizacionales • Cadenas revisadas de responsabilidad y rendición de cuentas • Planes de formación continua
4	<ul style="list-style-type: none"> • A.12. Desarrollar planes de mejora de calidad de datos 	<ul style="list-style-type: none"> • Planes de mejora de datos

En la etapa de **E (Explotación)** el objetivo es comprobar si se han alcanzado los objetivos del sistema de gobierno de datos (se han alcanzado los valores que se pretendían) mediante la re-ejecución de los procesos de negocio. Esta etapa aborda la tarea **Monitorizar** de ISO 38505-1/COBIT 2019. Por tanto, requiere haber diseñado y establecido los indicadores claves de procesos más representativos y útiles para dicha monitorización, así como los correspondientes cuadros de mandos. Esta etapa requiere la colaboración de roles de gobierno de datos que saben interpretar los valores de los indicadores para realimentar el siguiente ciclo KCAE, y roles de gestión de datos que proporcionan la información necesaria para computar los indicadores. Las tareas identificadas y los productos de trabajo se muestran en la Tabla 4.

4 Aplicación del modelo

El marco metodológico ha sido aplicado en diferentes organizaciones de distintos dominios. Por restricciones de espacio, referiremos brevemente una de estas aplicaciones en un servicio de una universidad española.

Tabla 4. Tareas y productos de trabajo típico de etapa E de Explotación

Mad.	Tareas	Productos Trabajo típicos
2	<ul style="list-style-type: none"> • E.01. Monitorizar la eficiencia de las tecnologías de información (DM.01) • E.02. Monitorizar la eficiencia de los procesos de negocio (DM.02) • E.03. Monitorizar la cobertura de los repositorios de metadatos (DM.03) • E.04. Monitorización de las medidas de seguridad (DM.04) • E.05. Monitorización de los niveles de calidad de datos para los activos de datos (DQM.01) • E.06. Monitorización del cumplimiento de las políticas (DG.01) 	<ul style="list-style-type: none"> • Informes de Tecnologías de Información • Informes de procesos de negocio • Informes de cobertura de repositorios de metadatos • Informes de calidad de datos.
3	<ul style="list-style-type: none"> • E.07. Monitorización del grado de cumplimiento de la estrategia de datos (DG.02) • E.08. Monitorización del ciclo de vida de los datos (DG.03) • E.09. Monitorización del desempeño de los recursos humanos (DG.04/DG.05) 	<ul style="list-style-type: none"> • Informe de cumplimiento de estrategia de datos. • Modelos de vida de los datos • Informes de desempeño de recursos humanos
4	<ul style="list-style-type: none"> • E.10. Monitorización de la eficiencia de las acciones correctoras de calidad de datos 	<ul style="list-style-type: none"> • Informes de calidad de datos

El objetivo era diseñar e implementar de forma iterativa e incremental el sistema de gobierno de datos, teniendo en cuenta las limitaciones, y la percepción de los beneficios que suponía para ellos. Tras una evaluación inicial del nivel de madurez usando MAMD, se determinó que no se alcanzaba el nivel 2, teniendo consolidado el nivel 1 (situación AS_IS). Tras un análisis de la viabilidad de la iniciativa y tras gestionar adecuadamente las expectativas, se determinó que el objetivo aceptable debería ser consolidar el nivel 2 (situación TO_BE). Se acordó una estrategia defensiva, con el valor intrínseco de los datos, y el valor de negocio de los datos a optimizar. Como parte de las etapas se concluyó que la situación de los componentes del sistema de gobierno de datos era la siguiente: (1) no se tenía un modelo de procesos DG/DQM/DM completamente descrito (aunque los procesos de DM estaban bastante avanzados), (2) no había roles específicos para el gobierno de datos, aunque sí para el gobierno de Tecnologías de Información, (3) no se habían establecido explícitamente políticas de datos ni se usaban marcos de referencia, estando todas las decisiones del uso y explotación de los datos basadas en la intuición de los dirigentes del servicio; (4) la única documentación explícita que existía sobre los datos era un diccionario de datos, que no siempre se utilizaba para la creación de consultas o servicios nuevos; (5) no hay una cultura explícita de datos, aunque se asume que son fundamentales para el desempeño del servicio; (6) como no se abordan roles ni responsabilidades de gobierno de datos, no se consideran importantes las habilidades ni competencias relacionadas, aunque tras las explicaciones pertinentes se consideró que tenerlas sería muy ventajoso, y finalmente (6), se tenían unos servicios/infraestructuras tecnológicas adecuadas que daban un soporte adecuado para la prestación de los servicios que proporcionaba, incluyendo niveles adecuados de seguridad para los datos. Durante la etapa A, se trabajó sobre todo en tres componentes del sistema de gobierno de datos: (1) la implantación de todos los

procesos hasta el nivel 2 de madurez de MAMD que mejor soporte podrían dar a las acciones de los procesos de negocio más críticos; basándose en esta implantación, se enriquecieron los procesos de negocio que se habían incluido en la auditoría de evaluación.; (2) se trabajó en identificar mejor alguno de los roles y responsabilidades asociadas para asignar trabajos específicos, mejorando así la infraestructura organizativa; y finalmente (3) se trabajó en identificar y documentar mejor los datos que se usaban en los procesos de negocio anteriormente mencionados, sobre todo el ciclo de vida de los datos y las reglas de negocio asociadas. Durante la etapa E, se constató que se había mejorado el valor intrínseco y de negocio de los datos, aunque se determinó que sería necesario mejorar el nivel de calidad de los datos. Con estos avances que consideró que el Sistema de Gobierno de Datos del servicio, había mejorado bastante, aunque quedaban algunas mejoras que se postpusieron para el siguiente ciclo KCAE.

5 Conclusiones y trabajos futuros

En este artículo se ha presentado un marco metodológico basado en diferentes estándares internacionales para el diseño, implantación y mantenimiento de sistemas de gobierno de datos. La implantación de este tipo de sistemas ayudará a la organización a garantizar que los datos que se usan en los procesos de negocio están alineados a la estrategia organizacional, con lo que su valor queda maximizado.

Como trabajos futuros contemplamos refinar el marco para aplicarlo en diferentes dominios. También pretendemos trazar el ciclo de vida de los diferentes productos de trabajo, sobre todo de los repositorios de conocimiento.

Agradecimientos

Esta publicación es parte de los proyectos de I+D+i PID2020-112540RB-C42, AETHER-UCLM (Una Aproximación Holística de Smart Data para el Análisis de Datos Guiado por el Contexto Centrada en la Calidad y la Seguridad), financiado por MCIN/AEI/10.13039/501100011033/, y ADAGIO (Alarcos' DATA Governance framework and systems generatIOn) (SBPLY/21/180501/000061), financiado por la Consejería de Educación, Cultura y Deportes de la Junta de Comunidades de Castilla-La Mancha) Esta investigación también ha sido cofinanciada por el Programa de Doctorado Industrial (Ref.: DIN2018-009705) del Ministerio de Ciencia, Innovación y Universidades.

Referencias

1. Aiken, P., Harbour, T.: *Data Strategy and the Enterprise Data Executive: Ensuring that Business and IT are in Synch in the Post-big Data Era*. Technics Publications (2017).
2. Allen, M., Cervo, D.: Chapter 6 - Establishing Multi-Domain Data Governance. In: Allen, M. and Cervo, D. (eds.) *Multi-Domain Master Data Management*. pp. 81–93 Morgan Kaufmann, Boston (2015). <https://doi.org/10.1016/B978-0-12-800835-5.00006-3>.

3. Al-Ruithe, M. et al.: A Conceptual Framework for Designing Data Governance for Cloud Computing. *Procedia Computer Science*. 94, 160–167 (2016).
<https://doi.org/10.1016/j.procs.2016.08.025>.
4. Carruthers, C., Jackson, P.: *The chief data officer's playbook*. Facet Publishing (2017).
5. DAMA: DAMA-DMBOK: data management body of knowledge. Technics Publications, LLC (2017).
6. Davenport, T., Harris, J.: *Competing on Analytics: Updated, with a New Introduction: The New Science of Winning*. Harvard Business Press (2017).
7. DMM: CMMI Institute - Data Management Maturity (DMM), <https://stage.cmmiinstitute.com/dmm>, last accessed 2021/03/31.
8. DQTeam: MAMD:Modelo Alarcos Mejora Datos, <https://mamd.dqteam.es>, last accessed 2022/04/11.
9. Engels, B.: Data Governance as the Enabler of the Data Economy. *Intereconomics*. 54, 4, 216–222 (2019). <https://doi.org/10.1007/s10272-019-0827-y>.
10. Eryurek, E. et al.: *Data Governance: The Definitive Guide*. O'Reilly Media, Inc. (2021).
11. EU: COM(2020) 66 Final, Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the regions a European Strategy For Data.
12. ISACA: COBIT 2019 Design guide: Designing an Information and Technology Governance Solution. , Schaumburg, IL. EE.UU (2018).
13. ISACA: COBIT 2019 Framework. Introduction and methodology. , Schaumburg, IL. EE.UU (2018).
14. ISACA: COBIT 2019 Implementation Guide: Implementing and Optimizing an Information and Technology Governance Solution. , Schaumburg, IL. EE.UU (2018).
15. ISO: ISO/IEC 38505-1:2017 Information technology — Governance of IT — Governance of data — Part 1: Application of ISO/IEC 38500 to the governance of data, <https://www.iso.org/standard/56639.html>, last accessed 2021/05/09.
16. ISO: ISO/IEC TR 38505-2:2018 Information technology — Governance of IT — Governance of data — Part 2: Implications of ISO/IEC 38505-1 for data management, <https://www.iso.org/standard/70911.html>, last accessed 2021/05/23.
17. Kazim, E., Koshiyama, A.: The interrelation between data and AI ethics in the context of impact assessments. *AI and Ethics*. (2020). <https://doi.org/10.1007/s43681-020-00029-w>.
18. Koltay, T.: Data governance, data literacy and the management of data quality. *IFLA Journal*. 42, 4, 303–312 (2016). <https://doi.org/10.1177/0340035216672238>.
19. Ladley, J.: *Data governance: How to design, deploy, and sustain an effective data governance program*. 2nd Edition. Academic Press (2019).
20. Laney, D.B.: *Infonomics: How to monetize, manage, and measure information as an asset for competitive advantage*. Routledge (2018).
21. OECD: *The Path to Becoming a Data-Driven Public Sector*. OECD Publishing, Paris (2019).
22. Osimo, D. et al.: *Data Strategies, Policies and Agenda*. 17 (2020).
23. Plotkin, D.: *Data stewardship: An actionable guide to effective data management and data governance*. Academic Press (2020).
24. Soares, S.: *Data governance tools: Evaluation criteria, Big Data governance, and alignment with enterprise data management*. Mc Press (2015).
25. Soares, S.: *The IBM data governance unified process: Driving business value with IBM software and best practices*. MC Press, LLC (2010).
26. Treder, M.: *The Chief Data Officer Management Handbook*. Apress, Berkeley, CA (2020). https://doi.org/10.1007/978-1-4842-6115-6_6.
27. Yébenes Serrano, J.R.: Marco para la construcción de sistemas de gobernanza de datos en entornos de industria 4.0. Framework for building data governance systems in industry 4.0 environments. (2022).