

La gobernabilidad de datos como apoyo en la gestión de datos de instituciones de educación superior

Data governance as support in the data management of institutions of higher education

Mariutsi Alexandra OSORIO Sanabria [1](#); Carlos Andrés GUERRERO Alarcón [2](#); Mayda Patricia GONZÁLEZ-ZABALA [3](#);

Recibido: 06/06/2017 • Aprobado: 05/07/2017

Contenido

- [1. Introducción](#)
 - [2. Fundamentos teóricos](#)
 - [3. Metodología](#)
 - [4. Resultados](#)
 - [5. Discusión](#)
 - [6. Conclusiones](#)
- [Referencias bibliográficas](#)

RESUMEN:

Los datos son considerados uno de los activos de información clave de los procesos de negocio de las organizaciones; por esta razón, existen marcos de trabajo para la gobernabilidad de datos que apoyan la gestión de información. Considerando que las Instituciones de Educación Superior (IES) requieren datos para la definición y evaluación de sus políticas y estrategias, se presenta el proceso de diseño de un plan de gobernabilidad de datos enfocado en la calidad, como apoyo a la toma de decisiones.

Palabras clave Gobernabilidad de datos, gestión de datos, calidad de datos

ABSTRACT:

Data are considered as one of the most key important assets of the business processes for all organizations; as a consequence, there are some frameworks to generate data governance supporting the information management process. Considering that permanently higher educational institutions (IES for its acronym in Spanish) are requesting data as inputs for defining and evaluating its policies and strategies, here it is shows the design process of a data governance plan, focused on quality as support for decisions making activities.

Keywords Data governance, data management, data quality

1. Introducción

Actualmente, las organizaciones impulsan su desarrollo en el conocimiento e información generada tanto en su interior como en el contexto. Para lograr su aprovechamiento requieren fuentes de datos confiables, coherentes y disponibles que soporten la toma de decisiones y el mejoramiento de sus capacidades operativas (Otto, Wende, Schmidt, & Osl, 2007). A raíz del aumento en la generación de los datos, dentro y fuera de las entidades, y con el aumento de las posibilidades de acceso a los mismos, cada vez más organizaciones se han dado cuenta de la importancia de gestionar y gobernar sus datos como un recurso a nivel estratégico que optimicen sus procesos (Fu, Wojak, Neagu, Ridley, & Kim, 2011; Niemi, 2011), de ahí que sean apreciados como otro tipo de activo clave que genera valor tal como los recursos financieros y humanos (DGI, 2015a; Kooper, Maes, & Lindgreen, 2011; Rifaie, Alhajj, & Ridley, 2009). Así mismo, han buscado mejorar la calidad y el seguimiento que hacen de sus datos con el fin de contar con nuevas formas de gestionarlos para generar información.

Las Instituciones de Educación Superior (IES) como organizaciones generadoras de conocimiento no son ajenas a esta situación dado que necesitan insumos y herramientas que apoyen la gestión de sus procesos misionales,

de manera que puedan responder a las demandas de su entorno (Osorio, 2016). Al igual que otras organizaciones, las IES soportan sus procesos administrativos, definen sus políticas y estrategias y basan sus decisiones en el análisis de sus datos, por tanto, se espera que estas instituciones utilicen las mejores prácticas para la gestión, intercambio, flexibilidad y control de calidad de sus datos.

No obstante, a pesar que las IES cuentan con plataformas tecnológicas que recopilan y almacenan sus datos, estos tienden a encontrarse dispersos, duplicados, incompletos o en diferentes formatos, pues son procesados en distintas aplicaciones. Esta situación hace que se puedan generar errores e inconsistencias en los reportes que son entregados a la comunidad interna y externa, quienes asumen la información recibida como cierta y veraz. Por esta razón, existen universidades que se han interesado en diseñar e implementar programas de *Gobernabilidad de Datos* (GD) que orienten a la alta dirección, sobre políticas, estándares, calidad y privacidad de datos, así como el desarrollo de Tecnologías de Información (TI) que permitan su uso y aprovechamiento, tales como aplicaciones de inteligencia de negocios, *data marts*, *data warehouse*, entre otras; buscando con ello mejorar la calidad de datos utilizados en la generación de informes, facilitar la rendición de cuentas, la seguridad y confidencialidad de los datos, el control y la distribución de información importante dentro de las universidades (New York University, 2015; Stanford University, 2015; Universidad of Cambrigde, 2014; University of Colorado, 2014; University of Michigan, 2016; University of Nevada Las Vegas, 2010; Young & McConkey, 2012).

Entre los aspectos para los cuales las IES han empezado a desarrollar planes de gestión de datos es para los relacionados con las actividades investigación, definiendo para ello el tratamiento que van a recibir los datos. La implementación de estos planes tiene como objetivos el garantizar que no se pierdan los datos, facilitar la correcta custodia desde su producción y permitir su preservación mediante documentación en el entorno tecnológico donde reposen (Preserva, 2013). De igual manera, se ha encontrado que uno de los factores que ha motivado el desarrollo de este tipo de planes es el hecho que los organismos financiadores de proyectos de investigación, establecen entre sus requerimientos la presentación de estos planes, tal como lo realiza la Unión Europea en su programa Horizonte 2020 (PAGODA, 2014).

Sin embargo, pese a los avances en materia de la formulación de iniciativas para gestionar los datos de las IES, un estudio realizado por (Osorio, 2016) evidenció que en la literatura científica, existen documentos orientados hacia la definición de marcos de trabajo de gobierno de datos que apoyen la gestión de los datos, pero no precisan casos de implementación de planes de Gobernabilidad de Datos (GD) con un enfoque hacia la calidad de los datos. En ese sentido, se identificó la oportunidad de desarrollar una investigación orientada a definir el diseño de un Plan de Gobernabilidad de Datos (PGD) enfocado en la calidad de datos para los procesos de investigación en las IES, de tal manera que se lograra un mayor aprovechamiento de los datos, la disminución de los tiempos y costos en la preparación y difusión de sus reportes. Cabe destacar, que esta investigación fue financiada por el programa de Talento Digital del Ministerio de Tecnologías de la Información y Comunicación (MINTIC) de Colombia y desarrollada en el marco de la Maestría en Tecnologías de la Información y Comunicación (TIC) de la Universidad Pontificia Bolivariana y aplicada en la Universidad Cooperativa de Colombia.

2. Fundamentos teóricos

2.1. Gestión, gobierno y calidad de datos

Teniendo en cuenta la importancia del valor de los datos y su incidencia en los diferentes aspectos humanos y en el desarrollo de las organizaciones, se llevó a cabo la revisión bibliográfica que permitió la identificación del alcance y la relación de los conceptos de gestión, gobierno y calidad de datos (Tabla 1).

Tabla 1 . Definición de gestión, gobernabilidad y calidad de datos

Concepto	Descripción
Gestión de datos	Ejercicio que orienta las actividades relacionadas con las funciones de los datos, implica el desarrollo, ejecución y supervisión de planes, políticas, programas y prácticas que controlan, protegen, entregan y mejoran el valor de los datos y la información como activos de las organizaciones (Mosley, M; Brackett, M; Earley , S; Henderson, 2009).
Gobernabilidad de datos	La disciplina que coordina personas, procesos y tecnología, con el fin de posicionar la información como un recurso de valor que satisfaga las necesidades de información de los interesados y contribuya en el cumplimiento de los objetivos de la organización, aumentando la confianza en la toma de decisiones a través de la calidad y seguridad de los datos (Osorio, 2016). Para (Khatri & Brown, 2010; Otto et al., 2007; Rifaie et al., 2009; Wende, 2007), la GD responde a preguntas sobre cómo se toman las decisiones relacionadas con los datos

	(dominios de decisión), quienes toman las decisiones (responsabilidades) y rinden cuentas, y cómo se miden y siguen los resultados de las decisiones, además se establecen directrices y normas para la gestión de la calidad de datos en toda la organización.
Calidad de datos	Es la percepción de los datos de acuerdo al propósito para el cual se utilizan en un contexto determinado y de cuan bien representan el mundo real (Alvarado, 2011). Este proceso requiere contar con actividades para la adquisición y estandarización de datos, identificación de fuentes confiables, integración, perfilamiento, identificación de objetos, localización y corrección de errores, y optimización de costos (Batini & Scannapieco, 2006).

Fuente: Elaboración propia

Para guiar el proceso de implementación de GD, IBM, ORACLE, el *Data Governance Institute* (DGI) y el DAMA han definido marcos de trabajo y funciones que benefician a las organizaciones en la generación de valor y en la coordinación e integración de las áreas de negocio (Adler, 2007; DGI, 2015b; IBM, 2007; Mosley, M; Brackett, M; Earley, S; Henderson, 2009; Oracle, 2011), así mismo, investigadores han propuesto modelos de GD centrados en roles, la identificación de responsabilidades en la toma de decisiones y acciones con el objeto de apoyar la gestión y calidad de los datos (Khatri & Brown, 2010; Thompson, Ravindran, & Nicosia, 2015; Weber et al., 2009), y hacia el diseño de procesos y estrategias que faciliten el control de los datos en aspectos o temas específicos (Al-Ruithe, Benkhelifa, & Hameed, 2016; Fu et al., 2011). De igual forma, se han llevado a cabo trabajos que resaltan la importancia de mejorar la gestión, la calidad, disponibilidad y seguridad de los datos (Fu et al., 2011; Guerrero & Londoño, 2016; Otto et al., 2007; Wende, 2007), otras propuestas se han derivado de marcos de gobierno de TI enlazándose con tendencias tales como la computación en la nube, *Big Data* o inteligencia de negocios (Al-Ruithe et al., 2016; Groß & Schill, 2012; Thomas, 2006; Weber et al., 2009).

Los modelos, marcos de trabajo y casos de estudio mencionados, proponen elementos, procedimientos, actividades y productos de la Gobernabilidad y calidad de datos, los cuales se integraron y compararon tal como se presentan en la Tabla 2.

Tabla 2 . Elementos, funciones o actividades de la gobernabilidad y calidad de datos

Marco o modelo	Gobernabilidad de datos			Calidad de datos
	Elementos	Procesos - Actividades	Productos o resultados	Procesos - Actividades
ORACLE	<ul style="list-style-type: none"> Política de Datos Interesados, responsabilidades, niveles y derechos de decisión, Estándares de datos 	<ul style="list-style-type: none"> Establecer, revisar, aprobar, supervisar las políticas y normas Establecer estrategias de datos empresariales 	<ul style="list-style-type: none"> Solución de problemas Proyectos y servicios de la gestión de datos Calidad de datos e información Reconocer el valor de los datos 	<ul style="list-style-type: none"> Establecer la dirección de calidad de los datos Monitorear la calidad de los datos Generar el informe de las iniciativas centradas en la calidad
Modelo madurez IBM	<ul style="list-style-type: none"> Estructura organizacional Política Representantes de calidad de datos 	<ul style="list-style-type: none"> Disciplinas soporte: arquitectura de datos, clasificación y metadatos, auditoría de información de Loggeo y reporte Disciplinas principales: calidad de datos, ciclo de vida del dato y seguridad y privacidad del dato 	<ul style="list-style-type: none"> Creación de valor Administración del riesgo de datos Cumplimiento 	
Guía DAMA-	<ul style="list-style-type: none"> Roles Políticas y estándares 	<ul style="list-style-type: none"> Estrategias de datos Gestión de problemas, comunicación 		<ul style="list-style-type: none"> Especificar los requerimientos de calidad Analizar el perfil de

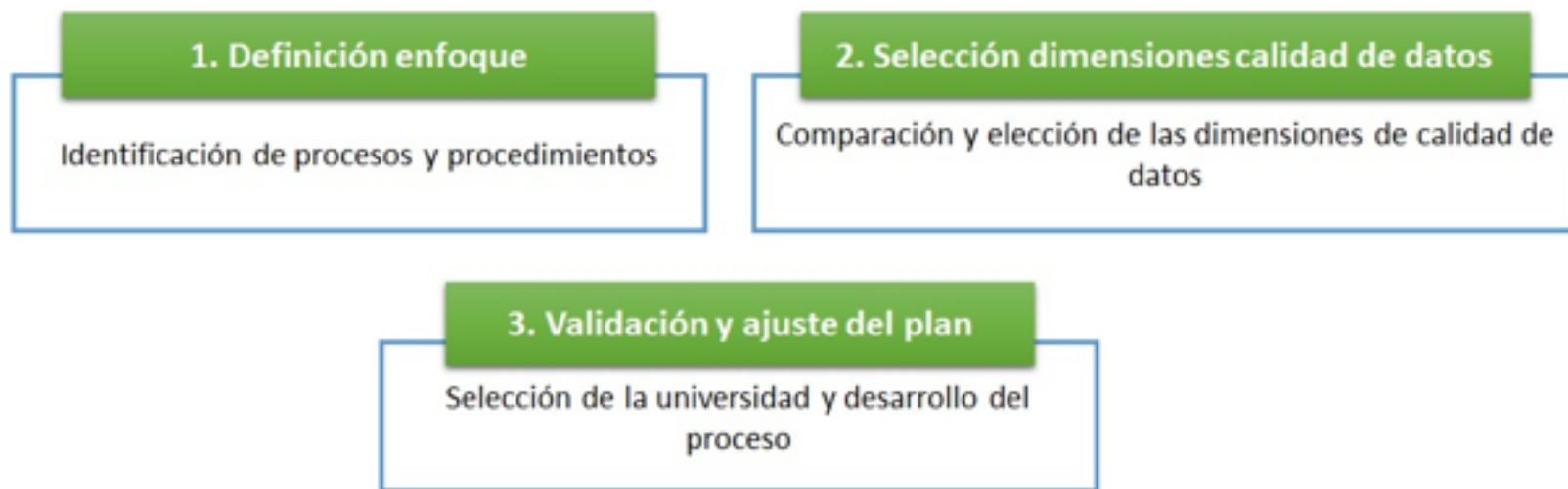
DMBOK	<ul style="list-style-type: none"> Proyectos y servicios 	<ul style="list-style-type: none"> Evaluación del activo de datos 	<ul style="list-style-type: none"> calidad Mejorar la calidad de los datos
Marco DGI	<ul style="list-style-type: none"> Personas Unidades de la organización Normas y reglas de compromiso para la gestión 	<ul style="list-style-type: none"> Procesos para gobernar los datos 	
Modelo de Madurez de Gestión de Datos		<ul style="list-style-type: none"> Gestión de gobierno Glosario de negocios Gestión de metadatos 	<ul style="list-style-type: none"> Definir estrategias Gestionar el ciclo de vida de dato Evaluar la calidad y realizar la limpieza de datos

Fuente: Elaboración propia basada en (Adler, 2007; DAMA-DMBOK, 2007; DGI, 2015a; IBM, 2007; Oracle, 2011; Serrano Alejo, 2015; Thomas, 2006)

3. Metodología

El proceso metodológico seguido para proponer el diseño del PGD contó con tres etapas basadas en los marcos de trabajo y modelos de implementación de gobierno y calidad de datos mencionados en la fundamentación teórica: 1) Definición del enfoque, 2) Selección de las dimensiones de calidad y 3) Validación y ajustes del plan (ver Figura 1).

Figura 1 . Etapas del proceso metodológico del trabajo



Fuente: Elaboración propia

3.1. Definición de enfoque

Oracle propone cinco enfoques para la GD desde una visión de la gestión empresarial de datos (Figura 2), teniendo en cuenta los conceptos de gobierno, gestión y calidad de datos presentados en la Tabla 1 , así como los elementos, procesos y actividades de la Tabla 2 se decidió que el diseño del PGD para una IES se realizara desde el enfoque de calidad de datos, entendiendo este proceso, como la idoneidad de los datos para servir a un propósito en un contexto dado, concepto relativo que toma diferentes significados de acuerdo al tipo de organización, las áreas específicas de aplicación y los usuarios (Fu et al., 2011; Ozmen-Ertekin & Ozbay, 2012).

Figura 2 . Enfoques y funciones claves dela GD bajo la visión de la gestión empresarial de datos



16	Precisión		X	X			
17	Relevancia / Pertinencia		X		X	X	X
18	Representación	X					
19	Reputación		X				
20	Seguridad		X				
21	Transparencia				X	X	
22	Unicidad	X					
22	Valor añadido		X				

Fuente: Elaboración propia basada en (Batini & Scannapieco, 2006; División de Estadística y Proyecciones Económicas CEPAL, 2003; "Global health elearning center", 2011; Instituto Nacional de Estadísticas de Chile, 2007; Lee, Strong, Kahn, & Wang, 2002; Valverde, 2014; Valverde, Marotta, & Vallespir, 2012)

Con base en necesidad de mejorar la recopilación y tratamiento de datos, y la comparación realizada, de las 22 dimensiones identificadas se seleccionaron tres para el enfoque establecido del PGD: exactitud, consistencia y completitud.

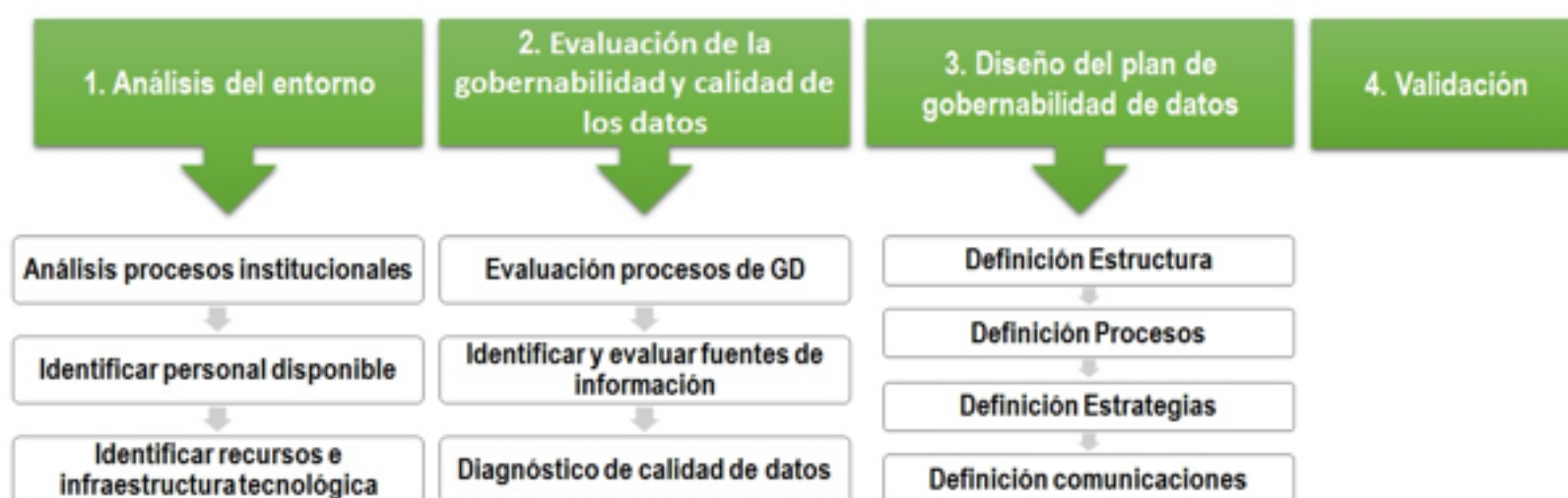
3.3. Validación y ajuste del plan

Para el desarrollo de la investigación se seleccionó como escenario y caso de estudio la Universidad Cooperativa de Colombia (UCC), universidad Multicampus (con presencia en 18 ciudades de Colombia y más de 50.000 estudiantes) que cuenta en su estructura organizativa con las direcciones de gestión tecnológica, investigación y planeación, lo cual facilitó la formulación, validación y ajuste del PGD.

4. Resultados

Se diseñó el PGD basado en las dimensiones de calidad de datos (exactitud, consistencia y completitud), siendo sus componentes: A) Estructura, B) Procesos, C) Estrategias y D) Comunicación, para su construcción se siguió el proceso de diseño ilustrado en la Figura 3 el cual contó con la validación de funcionarios de la Universidad tomada como caso de estudio (UCC). El PGD tiene como fin orientar los procesos, responsables y funciones de la gestión y calidad de los datos de investigación, para contar con un mayor aprovechamiento de los mismos en la preparación y difusión de reportes para la toma de decisiones. A continuación, se presenta cada componente del PGD diseñado y contextualizado para el caso de la UCC.

Figura 3 . Proceso del diseño del PGD



Fuente: Elaboración propia

4.1. Análisis del entorno

Para realizar la planificación es necesario comprender los objetivos y metas institucionales, e identificar de qué manera los activos de datos apoyan el cumplimiento de estos objetivos, para esto es necesario tener en cuenta

que la GD se impulsa por el valor de los datos, la rendición de cuentas, la medición del desempeño y la gestión de riesgos, aspectos que deben ser abordados en un proceso de gestión de datos (Rifaie et al., 2009). Para elaborar el PGD se recomienda llevar a cabo el análisis de los procesos institucionales en relación con la recopilación, almacenamiento y tratamiento de los datos. Es necesario identificar y evaluar las fuentes de información a través de un diagnóstico de la calidad de los datos. De igual forma se requiere identificar el personal disponible, los recursos e infraestructura tecnológica con que cuenta la institución, y los que puede vincular o adquirir para ser tenidos en cuenta en la implementación. En la Tabla 4 se presentan los elementos considerados en el análisis del entorno para el caso de la UCC.

Tabla 4 . Elementos para el análisis del entorno en la UCC

Proceso	UCC
Análisis de procesos institucionales	<p>Investigación: gestión de convocatorias, gestión planes de trabajo de los profesores investigadores, gestión de la producción académica y científica</p> <p>Tecnología: gestión de despliegue de las TIC y gestión del soporte de TIC</p>
Fuentes de información para diagnóstico de calidad de datos	<p>Bases de datos de gestión humana: profesores vinculados a la universidad y planes de trabajo.</p> <p>Bases de datos del Sistema Universitario de Investigación (SUI): grupos de investigación, investigadores, proyectos de investigación, y producción académica y científica)</p>
Personal	<p>Disponible: Arquitecta de procesos y de sistemas de información, Director nacional de investigación, Director de Planeación, Subdirectora de Autoevaluación, Especialista en gestión de información de investigación, Especialista en gestión de recursos financieros de investigación</p> <p>Por Vincular: Analista de calidad de datos,</p>
Recursos e infraestructura tecnológica	<ul style="list-style-type: none"> ● Intranet institucional ● Sistema de información laboral, Módulo de gestión humana (Human Content Management ● Sistema de información de gestión de planes de trabajo de profesores (CRONO) ● Bases de datos de la dirección de investigación, fondo editorial y comité nacional de productividad (Excel)

Fuente: Elaboración propia

4.2. Evaluación de la gobernabilidad y calidad de los datos

Se recomienda realizar la evaluación de la GD que se aplica en la entidad, en relación con la administración y riesgos de los datos asociados al proceso de investigación, esta evaluación permitirá identificar el impacto en los objetivos y metas institucionales, lo cual facilitará la definición de políticas, estándares y estrategias que faciliten la administración, el control y la calidad de los datos. Además, se recomienda evaluar las fuentes de información identificadas, realizando para ello el perfilamiento de datos que permita obtener el diagnóstico del estado de la calidad de los datos, donde se evidencien problemas y causas de la baja calidad, con el fin de identificar oportunidades de mejora.

En el caso de la UCC se diagnosticó la calidad de datos de dos fuentes de información de las seis identificadas (Sistema de información de planes de trabajo CRONO y bases de datos de la dirección de investigación), seleccionadas al evaluar cada fuente respecto a los indicadores de investigación de la Universidad (talento humano, producción académica, científica y tecnológica, estructuras de apoyo, proyectos e internacionalización de la investigación). Para identificar los problemas y causas de la baja calidad, se utilizaron las técnicas de perfilamiento de datos y vinculación de registros, lo cual permitió la definición de soluciones a los problemas de datos y estrategias de mejora (Tabla 5).

Tabla 5 . Problemas de datos, causas, propuestas de solución y estrategias de mejora

Problema	Causas	Propuesta de Solución	Estrategias de mejora
	<ul style="list-style-type: none"> ● Errores humanos, omisión en la 	<ul style="list-style-type: none"> ● Llevar a cabo un plan de verificación de los valores 	<ul style="list-style-type: none"> ● Definir e implementar un procedimiento que mida

Datos con valores nulos	digitación de campos relacionados con los registros de los proyectos de investigación o con la creación de los planes de trabajo de los profesores <ul style="list-style-type: none"> • Errores en la migración de datos antiguos 	faltantes con los soportes en físico con el fin de completar los registros <ul style="list-style-type: none"> • Definir el sistema CRONO como la fuente principal de ingreso de los datos de los proyectos de investigación aprobados. • Incluir validaciones de ingreso de datos al sistema CRONO, de manera que se reduzca el riesgo de crear registros incompletos. • Mejorar la integración del sistema CRONO con el sistema de gestión humana para vincular los datos de cédula y nombre del profesor de manera consistente • Corregir los parámetros de ingreso de campos clave. 	periódicamente la calidad de los datos. <ul style="list-style-type: none"> • Caracterización de los procesos de investigación para el rediseño de flujos de información, acciones, responsables y la actualización de procedimientos y formatos de captura de información • Diseño de un módulo en el sistema CRONO, para el registro y seguimiento de los proyectos de investigación • Programa de acompañamiento a los investigadores para facilitar la administración de información de investigación e interacción con las plataformas tecnológicas internas y externas
Datos inconsistentes	<ul style="list-style-type: none"> • Falta de validaciones en el archivo en el ingreso de los datos • Errores ortográficos y tipográficos • Migraciones de datos 		

Fuente: Elaboración propia

4.3. Diseño del plan de gobernabilidad de datos

Con base en el análisis del entorno y las evaluaciones de GD y calidad de datos, se procedió a la elaboración del plan, es así como se recomienda que sea documentado con ayuda de alguna plantilla. En el plan se deberá declarar explícitamente su alcance y cómo se llevará a cabo cada actividad. Además, se debe describir los elementos para cada uno de los componentes de la propuesta (estructura, procesos y comunicación) los cuales deberán articularse con los procesos de investigación y de la unidad responsable de gestionar los recursos tecnológicos de la de la institución.

4.3.1. Estructura (funciones y responsabilidades)

En este componente se definió:

- La visión, metas y estrategias del programa de GD a implementar, así como las estrategias y objetivos de TI, identificando sus responsables.
- El equipo de trabajo con sus objetivos y responsabilidades, teniendo en cuenta su interacción con los procesos y sistemas de información de la institución. El personal podrá asumir los roles de acuerdo a sus áreas de decisión o por su participación en algún comité o consejo (*data steward*, propietario o productor de datos, comité directivo de la gestión de los datos o consejo de gobierno de datos).
- Las políticas y normas necesarias relacionadas con la creación, adquisición, integridad, seguridad, cumplimiento y calidad de los datos de acuerdo a las necesidades del proceso de investigación, así mismo, se establecen sus objetivos y responsables, quienes deberán comunicarlas y realizar seguimiento de su aplicación y resultados. En la Tabla 6 se presentan las políticas de datos y responsables establecidos para la UCC.

Tabla 6 . Política de datos UCC

Política	Dueño de la Política
Los datos de investigación como activos de la Universidad	Vicerrectoría Académica
Acceso fácil y rápido de los datos de investigación	Dirección Nacional de Investigación
Seguridad de datos	Dirección Nacional de TI
Integridad de los datos	Arquitecto de procesos de TI
Política de Calidad y limpieza de datos	Data steward vinculado a la Dirección de TI
Política de tratamiento de información y protección de datos personales	Dirección de TI Dirección Nacional de Investigación

4.3.2. Procesos (procedimientos)

Como parte del plan se deben incluir los procesos y procedimientos que permitan gestionar eficazmente los datos almacenados en las fuentes de información institucionales garantizando su calidad. Los procesos propuestos se orientan hacia:

- **Supervisión y el seguimiento:** se recomienda definir indicadores que midan el desempeño de conjuntos específicos de datos, además se determina la métrica, meta, objetivo y responsable. En la UCC, se definieron los indicadores reportados en la Tabla 7 .

Tabla 7 . Indicadores de seguimiento del desempeño de un conjunto de datos

Métricas	Meta y Objetivo Métricas
Cantidad de datos duplicados y de datos faltantes	Monitorear que se está siguiendo la política de calidad y limpieza de datos
Recursos necesarios para la corrección de datos	Definir la asignación de recursos necesarios para la corrección de datos
Reducción del esfuerzo y recursos para resolver los problemas de datos	Mejorar las prácticas y procedimientos de la gestión de datos
Disminución de riesgo y costos por incumplimientos en el reporte de información a entidades externas	Cumplir con la entrega de informes confiables solicitados por entidades externas, de acuerdo a los requerimientos y tiempos establecidos
Número de problemas de calidad de datos recibidos y resueltos por el comité de dirección de gobernabilidad de datos	Evaluación de la gestión y eficiencia del comité de dirección de gobernabilidad de datos
Número de problemas de calidad de datos recibidos y resueltos por el consejo de gobernabilidad de datos	Evaluación de la gestión y eficiencia del consejo de gobernabilidad de datos
% de los costos operativos reducidos después de la aplicación del programa de gobernabilidad de datos	Aumentar la eficiencia del programa de gobernabilidad de datos
Número de veces que los datos críticos de investigación son cambiados o borrados de manera no autorizada	Garantizar la seguridad y protección de los datos
Reducción de los recursos necesarios para sincronizar los datos de múltiples fuentes	Asegurar la integración y disponibilidad de los datos de investigación

Fuente: Elaboración propia

- **Resolución de problemas:** Se recomienda definir los temas que podrían ser tratados por cada grupo equipo de trabajo, de acuerdo a su nivel de decisión, así como las reglas para los pasos de este procedimiento. En la Tabla 8 se presentan los temas clasificados en la UCC por nivel de decisión.

Tabla 8 . Temas a resolver por nivel de decisión

Nivel Estratégico	<ul style="list-style-type: none"> • Inversiones y asignación de recursos para la gestión de los datos • Incumplimiento del propósito de los datos
Nivel Táctico	<ul style="list-style-type: none"> • Las definiciones de los datos • Preocupaciones del uso e impacto de los datos
Nivel Operativo	<ul style="list-style-type: none"> • Calidad de los datos • Limpieza de los datos • Integridad de los datos

- **Auditoría y Monitoreo de datos:** Se recomienda definir el procedimiento para llevar a cabo un proceso una auditoría anual orientada al seguimiento de los datos, que permita revisar y mejorar las prácticas, los procesos y procedimientos que se desarrollan en el marco de la gestión y tratamiento de los datos dentro de la IES. Además, se debe establecer el conjunto de datos, herramientas y frecuencia del monitoreo de calidad de los datos.

4.3.3. Estrategias (proyectos y servicios)

Se recomienda definir estrategias, proyectos y servicios que apoyen la GD, para el caso de estudio en la Tabla 9 se presentan las iniciativas definidas en el PGD de la UCC:

Tabla 9 . Estrategias, proyectos y servicios definidos para el PGD de la UCC

Nº	Estrategias	Proyectos y servicios
1	Definir procesos, actividades y responsabilidades para la implementación y gestión del PGDI	<ul style="list-style-type: none"> • Implementación y seguimiento de las actividades definidas en el plan • Documentar las mediciones y evaluaciones de la calidad de los datos • Difundir, promover e implementar la política de tratamiento de información y protección de datos personales en la UCC
2	Establecer el plan de comunicación del GD y los mecanismos que permitan el trabajo colaborativo entre el equipo de trabajo	
3	Diseñar un ambiente de trabajo colaborativo para la implementación del PGDI	Programa de educación en gobernabilidad y calidad de datos con apoyo de herramientas tecnológicas, cursos y talleres en línea
4	La construcción de herramientas tecnológicas (Dashboard, Data Mart, BI, Big Data) que aporten en el análisis y explotación de los datos de investigación para la toma de decisiones institucionales	<ul style="list-style-type: none"> • Implementar el Sistema de Información de investigación • Implementar un dashboard y un Data Mart

Fuente: Elaboración propia

4.3.4. Procesos de Comunicación

Se recomienda definir los métodos y responsables del proceso de comunicación y cualificación entre la comunidad universitaria y grupos de interés, de acuerdo con los tipos de cambios que se generen o afecten la gestión de los datos de investigación, los cuales se pueden derivar por actualizaciones o nuevas políticas de datos, roles, procedimientos, reportes, resultados y evaluaciones. Para la UCC se definieron métodos y responsables del proceso de comunicación, a través de los cuales se informará al equipo de trabajo los cambios presentados en el PGD (Tabla 10).

Tabla 10 . Métodos y responsables de comunicación del PGD

Tipo de cambio	Método de comunicación	Responsabilidad
Proceso de Cambio	Correo electrónico a los grupos de interés	Data steward
Nueva política de datos de investigación	<ul style="list-style-type: none"> • Correo electrónico para la comunidad universitaria • Difusión medios institucionales 	Miembros del consejo de gobernabilidad de datos
Nuevos Roles	Anuncios informativos breves en reuniones de comité	Miembros del comité directivo de datos
Nuevos procedimientos	<ul style="list-style-type: none"> • Correo electrónico a los grupos de interés • Anuncios informativos breves en reuniones de comité 	Miembros del comité directivo de datos
Reportes de la evaluación de la calidad de los datos	Anuncios informativos breves en reuniones de comité y consejo	Miembros del comité directivo de datos

4.4. Validación

El PGD requiere la validación de parte de funcionarios adscritos a las unidades que lideran el proceso de investigación y de gestión de TI de la Institución, de manera que puedan liderar y promover su implementación en la institución. En la UCC, se validó la propuesta del PGD ante responsables y funcionarios de las Direcciones de gestión tecnológica, Investigación y planeación (Tabla 11).

Tabla 11 . Funcionarios que validaron la propuesta de PGD de la UCC

Dirección	Funcionarios de la UCC
Dirección nacional de gestión tecnológica	Arquitecta de procesos y de sistemas de información
Dirección nacional de investigación	<ul style="list-style-type: none"> • Director Nacional de Investigación • Jefes nacionales de investigación en posgrados (Facultades de Ingeniería y Medicina) • Especialista en gestión de información de investigación • Especialista en gestión de recursos financieros de investigación
Dirección nacional de planeación	<ul style="list-style-type: none"> • Director de Planeación • Subdirectora de Autoevaluación

Fuente: Elaboración propia

5. Discusión

Actualmente, las organizaciones necesitan contar con la gestión eficaz y eficiente de sus datos, con el fin de aprovechar nuevos modelos de negocio y mejorar sus capacidades operativas (Otto et al., 2007). El creciente interés en el tema de GD se debe a la importancia y ventajas del uso de los datos dentro y fuera de las organizaciones, dado que la GD brinda elementos para la mejora de la calidad de los datos almacenados, la reducción de costos y tiempo en la interpretación de información, así como en la construcción, difusión y entrega de reportes; además, aporta al acceso de datos precisos y al análisis de altos volúmenes por medio de herramientas de minería de datos, BI y Big Data (Osorio, 2016). Para lograrlo es necesario superar algunas barreras tales como la falta de políticas, procesos y funciones institucionales orientadas hacia la gestión de los datos, así como la falta de recursos, financiación, tecnología, habilidades, experiencia y conocimiento especializado de parte de los distintos actores que participan en este proceso (Al-Ruithe et al., 2016; Groß & Schill, 2012).

Se ha presentado el proceso para diseñar un PGD enfocado hacia la calidad de datos, que puede ser utilizado por IES como primer paso hacia la consolidación de un programa de GD, con el interés que cada vez más los datos sean tratados como un activo clave de la organización. En este diseño se han considerado cuatro etapas y cuatro componentes, a los cuales se les describen sus elementos con el fin de facilitar su aplicación. Si bien el plan se orientó hacia la calidad de datos por la importancia de contar con información confiable y oportuna, es posible que pueda tener otros enfoques lo cual dependerá de las necesidades de las instituciones, siendo el propósito que se pueda contribuir en intercambio de información y generación de conocimiento.

Esta propuesta fue validada con la implementación en la Universidad Cooperativa de Colombia, IES de carácter privado, la cual no cuenta con un programa de GD, pero que demandaba mejorar sus procesos de gestión de datos para la generación de indicadores e informes de investigación que permitieran el seguimiento de los objetivos de su plan estratégico nacional y atender las demandas de información de entidades externas. En dicha institución se siguió el procedimiento presentado en este trabajo para la construcción y puesta en marcha de algunos procedimientos, estrategias y proyectos establecidos en el PGD, permitiendo que la institución mejorará sus procedimientos de captura, limpieza y reporte de datos requeridos para la medición de sus capacidades en investigación, abordando los problemas de calidad de datos identificados. Cabe mencionar que a través de esta validación se realimentó la propuesta de este trabajo, al contar con los aportes de los funcionarios de la Universidad que participaron en su implementación.

Con base en el escenario de prueba de la propuesta, se identificó la importancia de GD para mejorar la calidad e integración de las fuentes de datos, considerando que las IES disponen de varios sistemas de información que almacenan sus datos en diferentes formas y formatos, lo cual hace que el seguimiento de la procedencia de los

datos y su combinación se convierta en una tarea compleja que demanda la coordinación y cooperación de las instancias que los administran, lo cual es requerido para la gestión de información.

De otra parte, la calidad de datos en las fuentes o repositorios, generalmente muestra la falta de aplicación de procedimientos que garanticen la completitud, consistencia y exactitud de los datos, así como su integración, genera costos en los procesos de construcción de reportes. Contar con estos atributos permite realizar una adecuada evaluación de la calidad de los datos, así como que estos sean compartidos con facilidad, en formatos compatibles con normas claramente definidas. Pese a esto, hay un factor que no se puede controlar, y es que estas evaluaciones se basan en la experiencia humana, y aunque existen técnicas que apoyan este proceso de manera automática, se hace necesario que los resultados sean analizados por expertos en el tema. Es así como se requiere contar con mecanismos y sistemas de información que integren los datos bajo criterios de calidad, de manera que esta pueda ser percibida sin tener que realizar procesos reiterativos de limpieza de los datos.

Se debe tener en cuenta las limitaciones de la propuesta, si bien se diseñó el PGD para IES que no cuentan con procesos formales de GD, no todas las instituciones cuentan con procesos definidos de gobierno de TI, lo cual hace que este tipo de ejercicios demande personal técnico y con experiencia, así como otros recursos no programados, siendo uno de los factores que puede incidir en la no implementación de lo establecido en el plan. Se espera que trabajos futuros orienten procesos que garanticen la seguridad y privacidad de los datos. Además, se espera que se deriven investigaciones posteriores que aborden temas relacionados con las métricas y mediciones de calidad.

6. Conclusiones

Con este trabajo, se pretende principalmente reflexionar sobre la importancia de la GD en las IES, así como contribuir con elementos que faciliten su implementación como parte del fortalecimiento de las capacidades de análisis y planificación en las instituciones. Se espera que esta propuesta, sea un primer paso que oriente la implementación de un programa de GD en una IES y contrarreste los actuales problemas de calidad de datos, de manera que posteriormente se construya un marco de referencia de GD para IES alineado con la gobernabilidad de Tecnologías de Información (TI).

El diseño de un PGD es un primer paso hacia la implementación de un programa de GD en cualquier entidad, este proceso requiere el consenso de las diferentes áreas que intervienen para que posteriormente se sigan las políticas, lineamientos y se lleven a cabo los mismos procedimientos que se han establecido, con el fin de garantizar la seguridad, la privacidad y la calidad de los datos, además de reducir los costos de la gestión de datos y los tiempos de entrega y análisis de información. El éxito de la implementación de un PGD, es posible si se cuenta responsabilidades claramente definidas, con procesos, procedimientos y herramientas apoyadas en TI que faciliten y guíen su gestión, así como que se cuente con el apoyo a nivel directivo y la disposición de los actores involucrados (investigadores, entidades financiadoras, participantes, IES, beneficiarios, etc.).

En la revisión de la literatura, se evidenció que la GD es objeto de estudio a nivel internacional y como proceso ha sido implementado en organizaciones que manejan grandes volúmenes de datos, pese a esto, en Colombia y en especial en el contexto universitario, la temática no es abordada con el mismo interés, lo cual incide en que las Direcciones de TI, no hayan adoptado e implementado estrategias, que promuevan la gestión de datos y con ella su gobernabilidad, situación que se presenta en la Universidad Cooperativa de Colombia. Por lo anterior, se recomienda que las instituciones dediquen esfuerzos que permitan crear una cultura donde los datos, la información y el conocimiento puedan ser gestionados y así contribuir al aprendizaje individual y colectivo entre los involucrados.

Finalmente, este trabajo aporta al área de gobierno de TI, en especial a las direcciones de tecnología de las IES, como guía para la configuración de los procesos que faciliten la implementación de un programa de GD centrado en la calidad de los datos. Como trabajo futuro, un grupo de investigación se encuentra desarrollando un marco conceptual para diseñar programas de GD, el cual se espera sea validado en otros escenarios por medio de estudios de caso, teniendo en cuenta que los resultados de la primera validación mostraron que es un recurso útil para estructurar y coordinar las responsabilidades del proceso de calidad, gestión y GD.

Referencias bibliográficas

- Adler, S. (2007). A Flash-light into the Dark: The Starters Guide to Data Governance. Recuperado a partir de http://www-07.ibm.com/events/au/iod/downloads/IPS6_Data-Governance.pdf
- Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2016). A conceptual framework for designing data governance for cloud computing. *Procedia Computer Science*, 94(MobiSPC), 160–167. <https://doi.org/10.1016/j.procs.2016.08.025>
- Alvarado, J. V. (2011). Calidad de datos en las organizaciones.
- Batini, C., & Scannapieco, M. (2006). *Data Quality Concepts, Methodologies and Techniques* (Springer). Springer-Verlag Berlin Heidelberg. <https://doi.org/10.1007/3-540-33173-5>

DAMA-DMBOK. (2007). Entidad de Conocimiento para la Gestión de Datos (DAMA-DMBOK ©) Marco de Trabajo Funcional.

DGI. (2015a). The Data Governance Institute. Recuperado 1 de enero de 2015, a partir de <http://www.datagovernance.com/the-dgi-framework/>

DGI. (2015b). The Data Governance Institute. Recuperado 9 de mayo de 2015, a partir de <http://www.datagovernance.com/dgi-data-governance-framework/>

División de Estadística y Proyecciones Económicas CEPAL. (2003). *Estudios estadísticos y prospectivos*. Santiago de Chile. Recuperado a partir de http://repositorio.cepal.org/bitstream/handle/11362/4725/1/S0311768_es.pdf

Fu, X., Wojak, A., Neagu, D., Ridley, M., & Kim, T. (2011). Data governance in predictive toxicology: A review. *Journal of Cheminformatics*, 3(7), 1–16. <https://doi.org/10.1186/1758-2946-3-24>

Global health elearning center. (2011). Recuperado 15 de junio de 2014, a partir de <https://www.globalhealthlearning.org/course/data-use-program-managers>

Groß, S., & Schill, A. (2012). Towards User Centric Data Governance and Control in the Cloud. *Lecture Notes in Computer Science*, 7039, 132–144. https://doi.org/10.1007/978-3-642-27585-2_11

Guerrero, C. A., & Londoño, J. M. (2016). Revisión de la Problemática de la Calidad del Software para el Desarrollo de Aplicaciones de Computación en la Nube. *Información tecnológica*, 27(3), 61–80. <https://doi.org/10.4067/S0718-07642016000300007>

IBM. (2007). *The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance. Governance An International Journal Of Policy And Administration*.

InfoTech. (2015). Info Tech Research Group. Recuperado 8 de febrero de 2015, a partir de <https://www.infotech.com/research/data-governance-implementation-plan-template>

Instituto Nacional de Estadísticas de Chile. (2007). *Dimensiones de la calidad según OECD y EUROSTAT. Reencuentro* (Vol. 51). Recuperado a partir de http://www.ine.cl/canales/menu/publicaciones/estudios_y_documentos/estudios/dimensionesdecalidad_4.pdf

Khatri, V., & Brown, C. V. (2010). Designing data governance. *Communications of the ACM*, 53(1), 148. <https://doi.org/10.1145/1629175.1629210>

Kooper, M. N., Maes, R., & Lindgreen, E. E. O. R. (2011). On the governance of information: Introducing a new concept of governance to support the management of information. *International Journal of Information Management*, 31(3), 195–200. <https://doi.org/10.1016/j.ijinfomgt.2010.05.009>

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information and Management*, 40(2), 133–146. [https://doi.org/10.1016/S0378-7206\(02\)00043-5](https://doi.org/10.1016/S0378-7206(02)00043-5)

Marotta, A. (2013). Dimensiones de Calidad. Montevideo. Recuperado a partir de [https://www.fing.edu.uy/inco/cursos/caldatos/Transparencias/2-Dimensiones de calidad.pdf](https://www.fing.edu.uy/inco/cursos/caldatos/Transparencias/2-Dimensiones%20de%20calidad.pdf)

Mosley, M; Brackett, M; Earley , S; Henderson, D. (2009). The DAMA Guide to the Data Management Body of Knowledge. DAMA.

New York University. (2015). New York University. Recuperado 16 de febrero de 2015, a partir de <http://www.nyu.edu/employees/resources-and-services/administrative-services/university-data-warehouse-plus/data-management.html>

Niemi, E. (2011). *Designing a Data Governance Framework*. Recuperado a partir de https://www.researchgate.net/publication/257194534_Working_Paper_Designing_a_Data_Governance_Framework

Oracle. (2011). *Enterprise Information Management: Best Practices in Data Governance. An Oracle White Paper in Enterprise Architecture*. Recuperado a partir de <http://www.oracle.com/technetwork/articles/entarch/oea-best-practices-data-gov-400760.pdf>

Osorio, M. (2016). *Diseño de un plan de gobernabilidad de datos de investigación para la Universidad Cooperativa de Colombia*. Universidad Pontificia Bolivariana.

Otto, B., Wende, K., Schmidt, A., & Osl, P. (2007). Towards a Framework for Corporate Data Quality Management Towards a Framework for Corporate Data Quality Management. *ACIS 2007 Proceedings*.

Ozmen-Ertekin, D., & Ozbay, K. (2012). Dynamic data maintenance for quality data, quality research. *International Journal of Information Management*, 32(3), 282–293. <https://doi.org/10.1016/j.ijinfomgt.2011.11.003>

PAGODA. (2014). PAGODA. Recuperado 18 de octubre de 2014, a partir de <http://www.consorciomadrono.es/pagoda/>

Preserva. (2013). Preserva. Recuperado 15 de febrero de 2015, a partir de <http://bd.ub.edu/preservadigital/node/180>

Rifaie, M., Alhajj, R., & Ridley, M. (2009). Data governance strategy: A key issue in building enterprise data warehouse. *iiWAS2009 - The 11th International Conference on Information Integration and Web-based Applications and Services*, 587–591. <https://doi.org/10.1145/1806338.1806449>

Serrano Alejo, S. (2015). *Herramienta de soporte para la auditoría de la madurez organizacional de la gestión de datos*. Escuela Superior de Informática. Universidad de Castilla - La Mancha. Recuperado a partir de https://ruidera.uclm.es/xmlui/bitstream/handle/10578/5085/TFG_Sergio_Serrano.pdf?sequence=1&isAllowed=y

Stanford University. (2015). *Getting Started with Stanford Data Governance*. Recuperado a partir de Getting Started with Stanford Data Governance

Thomas, G. (2006). The DGI data governance framework. Recuperado a partir de <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+DGI+Data+Governance+Framework#0>

Thompson, N., Ravindran, R., & Nicosia, S. (2015). Government data does not mean data governance: Lessons learned from a public sector application audit. *Government Information Quarterly*, 32(3), 316–322. <https://doi.org/10.1016/j.giq.2015.05.001>

Universidad of Cambrigde. (2014). University of Cambridge. Recuperado 10 de octubre de 2014, a partir de <http://www.data.cam.ac.uk/>

University of Colorado. (2014). *Data Governance*. Recuperado a partir de <http://www.cu.edu/sites/default/files/6010.pdf>

University of Michigan. (2016). Information technology stratetyc plan. Recuperado a partir de http://cio.umich.edu/system/files/IT_Strategic_Plan_-_2016.pdf

University of Nevada Las Vegas. (2010). *UNLV_Data_Governance_Policy.pdf*. Recuperado a partir de https://ir.unlv.edu/IAP/Files/UNLV_Data_Governance_Policy.aspx

Valverde, C. (2014). *Calidad de Datos en Experimentos de Ingeniería de Software*. Universidad de la República. Recuperado a partir de https://www.fing.edu.uy/sites/default/files/biblio/22369/tesis_carolina_valverde_-_cd_en_exp_de_is.pdf

Valverde, C., Marotta, A., & Vallespir, D. (2012). Análisis de la calidad de datos en experimentos en ingeniería de software. *XVIII Congreso Argentino de Ciencias de la Computación*. Recuperado a partir de <http://sedici.unlp.edu.ar/handle/10915/23711>

Weber, K., Otto, B., Osterle, H., Schneider, B., McManus, J., Kooper, M., ... Boff, S. O. (2009). One Size Does Not Fit All — A Contingency Approach to Data Governance. *ACM Journal of Data and Information Quality*, 1(1), 4:1-4:27. <https://doi.org/10.1145/1515693.1515696>.http

Wende, K. (2007). A model for data governance—Organising accountabilities for data quality management. *18th Australasian Conference on Information Systems*, 417–425. Recuperado a partir de http://www.alexandria.unisg.ch/export/DL/Kristin_Wende/204799.pdf

Young, A., & McConkey, K. (2012). Data governance and data quality – is it on your agenda? *Journal of Institutional Research*, 17.1(November 2011), 69–77.

1. Estudiante Doctorado en Ingeniería. Magíster en TIC becada Cv. Talento Digital del MINTIC. Universidad Cooperativa de Colombia. Bucaramanga, Colombia. Mariutsi.osorio@upb.edu.co

2. Doctor(c) en Ingeniería. Universidad Santo Tomás. Tunja, Colombia. carlos.guerrero@usantoto.edu.co

3. Doctora en Ingeniería. Universidad del Magdalena. Santa Marta, Colombia. mpgonzalez@unimagdalena.edu.co

Revista ESPACIOS. ISSN 0798 1015
Vol. 38 (Nº 51) Año 2017

[Index]

[En caso de encontrar un error en esta página notificar a [webmaster](#)]

©2017. revistaESPACIOS.com • ®Derechos Reservados