

Mejora en la gestión de Calidad de la información.



ISMAEL CABALLERO

Grupo de Investigación ALARCOS,
Escuela Superior de Informática,
(UCLM).
Paseo de la Universidad 4, 13071,
Ciudad Real (España)
{Ismael.Caballero,
[Mario.Piattini}@uclm.es](mailto:Mario.Piattini@uclm.es)
MiguelA.Blanco1@alu.uclm.es
PAIS (ESPAÑA)

MIGUEL ÁNGEL BLANCO

Grupo de Investigación ALARCOS,
Escuela Superior de Informática,
(UCLM).
Paseo de la Universidad 4, 13071,
Ciudad Real (España)
{Ismael.Caballero,
[Mario.Piattini}@uclm.es](mailto:Mario.Piattini@uclm.es)
MiguelA.Blanco1@alu.uclm.es
PAIS (ESPAÑA)

MARIO PIATTINI

Director, Grupo de Investigación
ALARCOS, Escuela Superior de
Informática, (UCLM).
Paseo de la Universidad 4, 13071,
Ciudad Real (España)
{Ismael.Caballero,
[Mario.Piattini}@uclm.es](mailto:Mario.Piattini@uclm.es)
MiguelA.Blanco1@alu.uclm.es
PAIS (ESPAÑA)

RESUMEN.

La calidad de la información es un factor cada día más decisivo para las organizaciones. Por eso, se han abierto muchas líneas de investigación para cubrir aspectos relacionados con esta creciente importancia. Muchos de estos esfuerzos han consistido en definir dimensiones o criterios de calidad para los datos y la información, o en definir metodologías de evaluación. Sin embargo, es necesario definir un marco de trabajo integrado que pueda guiar globalmente a las organizaciones en la evaluación y la mejora de la calidad de datos y de la información. En este artículo se propone un marco de trabajo que intenta llenar este vacío. Este marco de trabajo, basado en el concepto de Proceso de Gestión de Información, consta de dos componentes principales: un Modelo de Gestión de la Calidad de la Información estructurado en Niveles de Madurez (CALDEA) y una Metodología para la Evaluación y Mejora de la Gestión de la Calidad de la Información (EVAMECAL). La metodología permite la evaluación de los Procesos de Gestión de Información más importantes de la organización según los niveles de madurez propuestos por CALDEA, que a su vez sirve como guía para la mejora. También se muestra la aplicación a un caso práctico.

PALABRAS CLAVES

Aseguramiento de calidad de la Información, Proceso de gestión de Información, Niveles de Madurez para el proceso de gestión de información.

1 INTRODUCCIÓN

Actualmente la información es un factor decisivo para la economía en esta era de las Tecnologías de la Información [16], [20], [47] ya que es básica para la toma de decisiones tácticas, estratégicas u operativas [48], [54]. Investigadores y organizaciones han reconocido en este hecho la necesidad de considerar los datos y la información como uno de los valores mas importantes [27]. Una pobre calidad de datos y de

información tendrá un impacto negativo sobre la eficiencia global de las organizaciones [7], [40], [48]. Generalmente el aseguramiento de la calidad es un proceso complejo, en el que la diferencia entre costes y necesidades de calidad queda íntimamente vinculado por el contexto de aplicación y por los requisitos de la organización [6].

Las organizaciones necesitan una guía para los esfuerzos de evaluación y mejora de la calidad de la información de una forma integradora, coordinando y organizando los recursos para que se pueda alcanzar cierto grado de eficiencia en el tratamiento de sus datos teniendo en cuenta el conocimiento sobre la organización y los "procesos de fabricación de información" [56]. Estos esfuerzos comienzan con la

evaluación del estado actual de la calidad de los datos y de la información. Algunas investigaciones proporcionan varias medidas para la valoración de la calidad de la información y metodologías para la evaluación [13, 15, 27, 42, 45, 48, 58, 59], pero ninguna de ellas trata sobre cómo optimizar la calidad de la información coordinando esfuerzos o compromisos que alcancen a las organizaciones globalmente cubriendo aspectos teóricos y prácticos.

En [16] se establecen las siguientes cuatro características que debe proporcionar un buen marco de trabajo de calidad de la información:

- Un conjunto sistemático y conciso de criterios para evaluar la información.
- Un esquema para analizar y resolver problemas de calidad de información.
- Las bases para la gestión proactiva de calidad de la información a la comunidad investigadora.
- Un mapa conceptual que pueda usarse para estructurar varias teorías y fenómenos relacionados con la calidad de datos.

Nuestro objetivo es el desarrollo de un marco de trabajo integrador, que satisfaga las condiciones anteriores y cuyo propósito sea ayudar a las organizaciones a alcanzar sus objetivos de calidad de información mediante una gestión apropiada y organizada.

Nuestra propuesta está basada en dos principios: considerar la información como un producto [56] (lo que permite mirar la información desde un punto de vista ingenieril [4, 39] y poder aplicar los fundamentos de calidad de un proceso típico de producción [10, 11, 38]) y el concepto de Proceso Software dado por [19] (lo que permite estudiar mejor cómo se gestiona la fabricación). Estos dos principios sirven para definir el concepto de **Proceso de Gestión de Información (PGI)** como una visión unificada de ambos procesos, Producción de la Información y Gestión de la Calidad de la Información. Un PGI, por tanto, permite modelar cómo se produce la información y cómo se gestiona su calidad para una aplicación específica. Así para optimizar la calidad de la información debe evaluarse y optimizarse el PGI. La propuesta que se presenta da soporte a la evaluación y mejora del PGI. Aunque existen varios marcos de trabajo para la evaluación y la mejora de los procesos software, como CMM (Capability Maturity Model) [28, 46], CMMI (CMM Integration) [50], ISO 9001 [9], BootStrap [3] y ISO 15504 [35], ninguno de ellos se centran en la calidad de la información. Para evaluar un PGI específico es necesario compararlo con otro que se tome como referencia. La propuesta presenta ese modelo de referencia, especificando las actividades que se deben ejecutar para asegurar distintos objetivos de calidad de datos y de información en diferentes contextos. Para mejorar el PGI es necesario dar una serie de pasos, encaminados a añadir al PGI las actividades del modelo de referencia que no incluye. Como resultado, nuestra propuesta define dos componentes principales:

- Un Modelo de Gestión de Calidad de la Información basada en Niveles de Madurez, llamado **CALDEA**, donde se establecen distintas Áreas Clave de Procesos (ACP) que agrupan las actividades a las que se hacía referencia anteriormente. La razón de esta estructura en niveles es permitir conocer el avance en la mejora de la calidad de datos y de

información.

- Una Metodología para la Evaluación y Mejora de la Calidad de información, conocida como **EVAMECAL**, parecida a CBA-IPI [12], SCAMPI [53] o ISO 15504 [36], que consiste en un conjunto de pasos que proporciona una base para la medición y ejecución de la mejora de la calidad de datos y de la información a través de una gestión proactiva.

El resto de este artículo se estructura de la siguiente forma: La sección 2, resume brevemente el Modelo de Calidad de Información basado en Niveles de Madurez. La sección 3, describe la Metodología de la Evaluación y Mejora de la Calidad de la Información. La sección 4, muestra una aplicación práctica del marco de trabajo a un caso de estudio. Y la sección 5, detalla algunas conclusiones líneas de trabajo de investigación futuro.

2 CALDEA: UN MODELO DE CALIDAD DE DATOS Y DE INFORMACIÓN BASADO EN NIVELES DE MADUREZ.

CALDEA define cinco Niveles de Madurez de Gestión de la Calidad de la información para un PGI: Inicial, Definición, Integración, Gestión Cuantitativa y Optimizante. Los niveles se han estructurado teniendo en cuenta los distintos objetivos de calidad de la información y sus importancias relativas, proporcionando un conjunto sistemático y conciso de criterios que pueden ser usados para garantizar la calidad buscada [16]. CALDEA identifica las ACPs técnicas y de gestión que recogen objetivos específicos de calidad de información. Cada ACP contiene actividades, que se pueden ejecutar usando técnicas, prácticas y herramientas. La elección de las ACPs para cada nivel, y sus correspondientes actividades y tareas están basadas en dos pilares: las ACPs de CMMI [50] y las lecciones aprendidas como resultado de nuestras experiencias en las distintas iniciativas en las que hemos participado.

2.1 NIVEL INICIAL

Un PGI se encuentra en el Nivel Inicial cuando no se realiza ningún esfuerzo para alcanzar los objetivos de calidad de información.

2.2 NIVEL DE DEFINICIÓN

Un PGI está en este nivel cuando se ha definido y planificado. Esto implica identificar todos sus componentes y sus relaciones. Se tienen que satisfacer las siguientes ACPs:

- **Gestión de Fuentes de Datos y Destinos de Producto de Información (FS)**. Por las características de los datos, es necesario identificar y documentar las fuentes de datos y los objetivos de los datos de ERU-PGI para evitar problemas de redundancia y de formato de intercambio de los datos [43]. En [1,5] se discute estos temas y sugiere formas para tratar la información de múltiples fuentes. En almacenes de datos se deben usar herramientas como ETLs para unificar la semántica y los formatos de datos [13].
- **Gestión de Proyecto de Mantenimiento, Desarrollo o Adquisición de Bases de Datos o Almacenes de Datos (ADM)**. Los datos usados como materias primas deben ser recogidos y almacenados en BD apropiadas o en almacenes de datos. Para asegurar mejor la calidad de la

información, es mejor elaborar un proyecto para la gestión de la adquisición, desarrollo o mantenimiento de un SGBD o un almacén de datos, debiendo soportar los requisitos establecidos en ERU-PI, ERU-CI y ERU-PGI. Esta ACP permite incluir otras actividades como Aseguramiento de la Calidad de los Datos [37].

- **Gestión de la Calidad de la Información en Componentes PGI (GCI).** El uso de métricas para medir la eficiencia del PGI permite ayudar a mejorarlo. Es necesario identificar a partir de las ERU-CI tanto las dimensiones de calidad de la información (como se propone en [34] para el software) para cada uno de los componentes que se debe controlar [26, 27], como las métricas que mejor se adapten a cada una de estas dimensiones [15, 39, 47]. Para conseguir que nuestro marco sea lo más genérico posible, no se propone ninguna dimensión como obligatoria, ya que esto no es posible salvo para contextos concretos [47]. Como guía, suele utilizarse como estándar el conjunto de dimensiones de calidad de datos propuesto por [54]. Como ayuda en esta labor, se puede utilizar una metodología genérica como [32] o GQM descrita en [55]. Autores como [2, 5, 8] han propuesto métricas para la mejora de componentes específicos de un PGI. Para obtener medidas más fiables es interesante automatizar el proceso de medición como sugiere [24].

2.3 NIVEL DE INTEGRACIÓN.

Un PGI está en un Nivel de Integración cuando además de haber ejecutado el Nivel de Definición, se realizan esfuerzos para asegurar que el PGI cumpla con los requisitos, estándares y políticas de calidad de la información que marca la organización. Implica estandarizar lo aprendido en las diferentes experiencias de calidad de la información, a través de estándares y políticas de calidad de la información para evitar errores previos y permitir un mejor trabajo en el futuro. Las ACPs que se tienen que cumplir son:

- **Validación y Verificación de Componentes del PGI y Productos de Información (VV).** Tanto los productos de información como los componentes del PGI tienen que ser verificados y validados para corregir defectos y diferencias con las ERUs. Se podría usar las inspecciones de software [17, 22], adaptadas a los aspectos de calidad. Una metodología más específica que se puede usar es, data testing propuesta por [41], extendida al PGI. Se podría diseñar y redactar un plan de pruebas para coordinar los esfuerzos relativos a la validación y a la verificación [30].
- **Gestión del Impacto de la Calidad de Información y de los Riesgos (GIR).** Es necesario determinar el impacto sobre la organización de una pobre calidad de la información en el PGI y acotar qué riesgos se producirán si se asume [13]. [21] propone una metodología que se puede adaptar a aspectos de calidad de la información para recoger y documentar todos los riesgos.
- **Gestión de la Estandarización de la Calidad de**

la Información (GECI). Todas las lecciones aprendidas a través de la experiencia se deben recoger, documentar y transmitir a la base de conocimientos de la organización, Un ejemplo de proceso de estandarización podría ser el sexto proceso de TQDM [14].

- **Gestión de las Políticas de Calidad de la Información Organizativa (GPCIO).** Una forma de implementar todos los esfuerzos mencionados anteriormente, consiste en la definición de políticas de calidad de la información basadas en estándares previamente definidos, que afectan a la organización.

2.4 GESTIÓN CUANTITATIVA.

Un PGI se encuentra en este nivel cuando además de haberse consolidado el Nivel anterior, se han definido e implementado varios Planes de Medidas y los procedimientos de medidas se han automatizado. El objetivo de este nivel es obtener una conformidad cuantitativa automatizada de que el rendimiento del PGI en un periodo razonable de tiempo permanece fiel a los requisitos de calidad de datos y de información en términos de variación y la estabilidad según un conjunto fiable de medidas de dimensiones de calidad de la información para el PGI [18]. Las ACPs de este nivel son:

- **Gestión de Planes de Medición del PGI (GPM).** El objetivo de esta ACP es conseguir métricas que se pueden usar para comprobar la conformidad de la especificaciones [23, 43]. Como afirma [44], un plan para la medición de la calidad del software comienza con la decisión de tomar medidas e implica elegir "qué", "cuándo" y "cómo" medir, cómo representar esas medidas y "a quién". Autores como [13, 43] proponen el uso de diagramas de control como forma de representar los datos referentes al PGI. Otra forma de mostrar los resultados es el uso de diagramas de Kiviat como propone [29].
- **Gestión para la Automatización de Planes de Medida (GAPM).** Para conseguir un aumento de fiabilidad y repetibilidad de las medidas, algunos de los procedimientos de medidas y algoritmos (definidos en GM) deben ser automatizados [24].

2.5 NIVEL DE MEJORA.

Un PGI se dice que está en un Nivel Optimizante cuando además de gestionar cuantitativamente la obtención de medidas, éstas se usan para desarrollar un proceso continuo de mejora. Las dos siguientes ACPs deben satisfacerse:

- **Análisis Causal para la Gestión de Prevención de Defectos (GPD).** El Control Estadístico del Proceso (CEP) pueden aplicarse para detectar defectos de calidad de información e identificar sus causas. Las conclusiones obtenidas proporcionan una base para eliminar defectos detectados en los recursos afectados. [52] ofrece un marco de trabajo para prevenir los defectos.
- **Gestión de la Innovación y del Desarrollo de la Organización (GIDO).** Son las bases para el concepto de mejora continua. Similar al anterior ACP, los resultados se pueden usar para mejorar el PGI.

3 EVAMECAL: UNA METODOLOGÍA PARA LA GESTIÓN DEL ASEGURAMIENTO Y LA MEJORA DE LA CALIDAD DE LA INFORMACIÓN.

3.1 CONSIDERACIONES PREVIAS.

Antes de comenzar la presentación de la metodología es interesante explicar los términos en los que se va a llevar a cabo la evaluación y la mejora. Los diferentes estados para cada elemento del marco pueden ser clasificados como sigue:

- Cada Nivel de Madurez puede estar en uno de estos estados: {"Consolidado"/"No Consolidado"}.
- Cada ACP puede estar en uno de estos estados: {"Completamente Satisfecho", "Satisfecho", "Parcialmente Satisfecho", "No Satisfecho"}.

Cada actividad en las ACPs puede estar en uno de estos estados: {"Completamente Ejecutado", "Ejecutado", "Parcialmente Ejecutado", "No Ejecutado"}

Cada componente puede estar en uno de estos estados: {"Completamente Optimizado", "Optimizado", "Parcialmente Optimizado", "No Optimizado"}.

Para definir cuándo un elemento alcanza un estado concreto, una primera aproximación consiste en que cada organización defina sus propios criterios. Esta solución tiene dos problemas: en primer lugar, no es posible reflejar el grado de importancia que tiene cada elemento para la calidad global de la información; y en segundo lugar no sería posible comparar dos organizaciones, si cada una toma criterios diferentes.

Para solucionar el primer problema, se introdujo el concepto de Grado de Criticidad(GC) para cada elemento del modelo, definiendo así una guía de mejora: se debe optimizar antes los más críticos. Como complemento a este GC se definió para cada elemento el concepto de Valor de Calidad de la Información (VCI), como la media ponderada de la calificación que obtuviese cada subelemento en la valoración. Con el VCI se puede definir rangos de valores aceptables para cada uno de los estados en los que pueden estar los elementos. La tabla 1 muestra los distintos GCs del marco; estos valores han sido propuestos en base a nuestra experiencia.

Por otra parte, y teniendo cuatro estados para componentes, actividades y ACPs, se ha establecido (al estilo de [36]), valores para el VCI, estos son:

- Si 0 VCI 20, el estado es " No Optimizado/Ejecutado/Satisfecho".
- Si 21 VCI 60, el estado es "Parcialmente Optimizado/Ejecutado/Satisfecho"
- Si 61 VCI 90, el estado es "Optimizado/Ejecutado/Satisfecho",
- Si 91 VCI 100, el estado es "Completamente Optimizado/Ejecutado/Satisfecho".

Y para los niveles de madurez:

- Si 0 NM-VCI 90 el estado es "No Consolidado".
- Si 91 MN-VCI 100 y todas las ACP pertenecen como mínimo al estado "Satisfecho", su nivel es "Optimizado".

La validación teórica y empírica de estos GCs y de los rangos representa parte de nuestra futura línea de investigación en aras de una mayor aplicabilidad independientemente del contexto de la organización.

3.2 LA METODOLOGÍA DE EVALUACIÓN Y MEJORA

EVAMECAL tiene como principal objetivo la evaluación y mejora de un PGI de una organización dada. Sigue el modelo TQDM de Definir-Medir-Analizar-Mejorar, le añade un nuevo paso para la estandarización de las lecciones aprendidas.

La definición de objetivos, los procesos de medida, los criterios de análisis y los planes de mejora se han establecidos en los términos de calidad de información dados por CALDEA. Brevemente resumido, EVAMECAL puede ser enunciado como sigue:

- **Medir el estado actual del nivel de madurez.** El objetivo es determinar cuál es el estado actual del PGI en términos de niveles de madurez y de los sus componentes en términos de sus VCI. Para ello, se realizan cuestionarios, y si se han definido métricas, se toman medidas en relación con los componentes del PGI, para calcular los VCI.
- **Definir objetivos de mejora en términos de niveles de madurez de calidad de información.** Una iniciativa de calidad de información debe definir un plan que cubra los objetivos a mejorar. Estos objetivos han de expresarse en términos de niveles de madurez o de incrementos de VCI.

GRADOS DE CRITICIDAD	
Nivel de Definición	
(GEAGI) Gestión de un Equipo para el Aseguramiento de Calidad de la Información.	10 %
(GP) Gestión de Proyecto para PGI	15 %
(GRU) Gestión de Requisitos de Usuarios.	25 %
(FS) Gestión de Fuentes de Datos y Destinos de Producto de Información	10 %
(ADM) Gestión de Proyecto de Mantenimiento, Desarrollo o Adquisición de Bases de Datos o Almacén de Datos.	25 %
(GCI) Gestión de la Calidad de la Información en Componentes PGI.	25%
Nivel de Integración	
(VV) Validación y Verificación de Componentes del PGI y Productos de Información	25%
(GIR) Gestión del Impacto de la Pobre Calidad de Información y de los Riesgos.	25%
(GECI) Gestión de la Estandarización de la Calidad de la Información	25%
(GPCIO) Gestión de las Políticas de Calidad de la Información Organizativa.	25%
Nivel de Gestión Cuantitativa	
(GPM) Gestión de Métricas del PGI	70 %
(GAPM) Gestión para la Automatización de Planes de Medida	30 %
Nivel de Mejora	
(GPD) Análisis Causal para la Gestión de Prevención de Defectos	50%
(GIDO) Gestión de la Innovación y del Desarrollo de la Organización	50%

Tabla 1. Grados de criticidad.

- **Desarrollar un plan para la mejora teniendo en cuenta los objetivos marcados.** Un plan establece qué acciones deben ser ejecutadas para obtener el

conjunto de mejoras definidos en la fase 2. También define cuándo, cómo y quién debe ejecutar estas acciones, por quién y qué recursos están implicados.

- **Comprobar la eficiencia del plan de mejora.** Para cotejar la eficiencia del plan de mejora es necesario volver a realizar la evaluación, y comprobar si se han alcanzado los objetivos de calidad de información deseados.

- **Estandarizar las lecciones aprendidas para evitar futuros problemas.** El Equipo de Aseguramiento de Calidad de Información debe estandarizar el conocimiento adquirido a través de experiencias para evitar futuros problemas y conseguir mejores resultados.

3.3 LOS CUESTIONARIOS

Para el proceso de valoración se han redactado un conjunto de cuestionarios. Este conjunto consta de cuatro clases diferentes de cuestionarios, con diferentes objetivos:

A. Delimitar y caracterizar la organización, con un total de quince preguntas.

B. Delimitar y caracterizar el PGI a ser evaluado, con un total de seis preguntas.

C. Evaluar el grado de consolidación de cada nivel de madurez, para los que se han desarrollado varias preguntas organizadas en bloques diferentes y selectivos.

Las preguntas están enfocadas hacia las ACPs, actividades, tareas, técnicas propuestas, herramientas y prácticas y productos desarrollados y requeridos. La idea de organizar el cuestionario en varios niveles de profundidad es para que las preguntas se realicen desde arriba hacia abajo solo si es necesario. De esta forma, el bloque del primer nivel de profundidad sirve para evaluar si las ACPs se satisfacen o no, evitando realizar otras preguntas que no son importantes para establecer los aspectos específicos para la ejecución de actividades específicas dentro de cada ACP. De esta manera, si todas las respuestas de las preguntas del primer nivel de profundidad son distintas a "No Satisfecho", entonces se realizarán las preguntas del nivel inmediatamente inferior, y así sucesivamente. En total se contestarían ciento noventa preguntas en el caso que todas las respuestas a las preguntas del primer y segundo nivel de profundidad fueran diferentes a "No Satisfecho". La respuesta a estas preguntas debe ser un número entre 0 y 100, para calcular el correspondiente VCI para cada componente, actividad, ACP y Niveles de Madurez. Tenemos que señalar que estamos trabajando para comprobar la validez y eficiencia de cada bloque y dentro de cada bloque sus preguntas. No se incluyen las preguntas debido a las restricciones de espacio.

NIVELES DE PROFUNDIDAD	1	2	3
Nivel de Madurez 2	12	18	82
Nivel de Madurez 3	8	7	24
Nivel de Madurez 4	3	2	11
Nivel de Madurez 5	4	2	11
Preguntas Totales por niveles de profundidad	28	31	131
Preguntas Totales por Cuestionario	190		

Tabla 2. Numero de preguntas por niveles de profundidad y de madurez.

D. Finalmente, para comprobar y comparar las respuestas de los cuestionarios previos, se ha redactado un último bloque de preguntas con respuestas en lenguaje textual y descriptivo.

4 APLICACIÓN PRÁCTICA DEL MARCO DE TRABAJO A UN CASO DE ESTUDIO

A fin de validar empíricamente el marco de trabajo, se ha aplicado a diferentes PGI de varias organizaciones. A continuación se presentan los resultados obtenidos cuando se aplicó a un PGI de una organización con experiencia en gestión de información. EVAMECAL fue aplicado siguiendo los pasos previamente descritos. El resultado de los cuestionarios se presenta en las siguientes secciones, 4.1, 4.2 y 4.3.

4.1 DESCRIPCIÓN DE LA COMPAÑÍA.

La principal actividad de la compañía es desarrollar software. Poseen un sólido conocimiento y formación en estándares de calidad de software e Ingeniería del Software. La compañía ha obtenido la certificación ISO 9000. Los servicios que ofrece son consultoría, desarrollos a medidas, formación, atención técnica, ventas de licencias, bases de datos y administración de almacenes de datos, proyectos de planificación de sistemas y migración a importantes SGBD. Con un total de ochenta y nueve empleados, los dieciocho del Departamento de Sistemas son los que organizan los recursos que dan soporte informático al resto de los departamentos.

4.2 CARACTERIZACIÓN DEL ESTUDIO DEL PGI.

Entre todos los PGI de la organización, el marco de trabajo se aplicó al Proceso de Gestión de Formación, que es responsabilidad del Departamento de Consultoría. El principal objetivo del PGI estudiado es la gestión de los datos relativos a la formación, consistente en satisfacer demandas internas y externas para la formación, eligiendo quién va a ser el formador, determinando qué recursos se van a usar y gestionando distintos aspectos de calidad de la formación. Este proceso está adecuadamente especificado y documentado en el manual de calidad de la empresa.

Existen formularios para recoger datos sobre cursos demandados y para las evaluaciones de la calidad de los ejercicios propuestos, materiales didácticos proporcionados, y para la capacidad de los profesores, instalaciones, asistencia y uso de recursos.

La organización utiliza para la gestión de estos temas una aplicación llamada SINCRO, desarrollada internamente. Uno de los empleados del Departamento de Consultoría, es el responsable de transcribir datos de los formularios a la aplicación y obtener la información que será transmitida a la persona adecuada.

4.3 EVALUACIÓN Y MEJORA DE PGI.

Es importante destacar que tras la evaluación se acordó un tiempo de seis meses para implementar las mejoras propuestas como resultado de la aplicación de EVAMECAL.

Se comenzó con la evaluación del Nivel Defindo. Se fueron proponiendo las preguntas del cuestionario al director del Departamento de Consultoría, quien las respondía y comentaba oportunamente. Las conclusiones extraídas fueron las siguientes:

- Gestión del Equipo para la Gestión de la Calidad de la Información (GEACDI). A pesar de que la organización cuenta con un equipo de

aseguramiento de calidad, no se habían definido ni los roles ni las responsabilidades respecto a la calidad de los datos y de la información. Por ello, esta ACP estaba en un nivel de "No Satisfecho". Para mejorar esta ACP, se propuso la creación de un EACDI, la identificación de roles y de responsabilidades. Tras el periodo pactado para implementar las mejoras se logró alcanzar el nivel de "Parcialmente Satisfecho".

- Gestión de Proyectos para PGI (GP). Por otro lado, tampoco se había formalizado la realización de un proyecto para este PGI, con lo que esta ACP estaba en un nivel de "No Satisfecho". Como propuesta de mejora se sugirió la formalización de un proyecto para el PGI de estudio, pero durante el tiempo de implementación de mejoras no se consiguió dedicar recursos a esta ACP, con lo que continuó en el nivel de "No Satisfecho".
- Gestión de Requisitos de Usuarios (GR). Los requisitos de usuarios se gestionan para los procedimientos de formación, aunque los requisitos de calidad de información no han sido tratados de ninguna manera. Como se tiene una infraestructura para dar soporte a esta ACP, y teniendo en cuenta las respuestas del interlocutor, se evaluó esta ACP con el nivel "Parcialmente Satisfecho". Para mejorar la gestión de requisitos de usuario se propuso una formalización de los requisitos de usuarios para el PGI y para la calidad de los datos y de la información. Durante el tiempo de mejoras, se consiguió recoger y documentar ambos tipos de requisitos, pero no se llegaron a implementar. La nueva evaluación dio como resultado que esta ACP estaba en el nivel de "Satisfecho".
- Gestión de las Fuentes de Datos y Destino de los Productos de Información (FS). Tanto las fuentes de los datos como los destinos de los productos están completamente identificados, aunque no existe una uniformidad en los formatos de los datos. Esto hacía que esta ACP estuviese en un nivel de "Parcialmente Satisfecho". Se propuso como mejora una unificación de los formatos de datos que entraban al sistema. Se consiguió parcialmente este objetivo, con lo que la evaluación subió hasta el nivel de "Satisfecho".
- Gestión de Proyecto de Mantenimiento, Desarrollo o Adquisición de Bases de Datos o Almacenes de Datos (ADM). La base de datos donde los datos son almacenados pertenece a la organización y ha sido adecuadamente modificada para dar soporte a la aplicación SINCRO. Los modelos de datos y de procedimientos fueron modificados y extendidos, pero sin embargo no observan ninguno de los aspectos de calidad de la información. La calificación obtenida en esta ACP fue la de "Parcialmente Satisfecho". Como mejora se propuso documentar convenientemente la base de datos en la que SINCRO almacenaba los datos y llevar un control de configuración de los cambios. Se consiguió hacer el control de configuración. La calificación subió al nivel de "Satisfecho".
- Gestión de la Calidad de la Información en Componentes PGI (GCI). Esta ACP no se ha conseguido porque la calidad de la información no

está contemplada como un objetivo en PGI. Inicialmente su calificación fue la de "No Satisfecho". Se recomendó elegir las dimensiones de calidad de datos y de información que podían afectar a cada uno de los componentes del PGI. Se identificaron algunas dimensiones de calidad para los valores de los datos y para el modelo relacional de la base de datos y se aplicaron algunas de las métricas. La calificación de este nivel subió a "Parcialmente Satisfecho".

Haciendo los cálculos correspondientes, se deduce que el PGI "Gestión de Formación" no ha consolidado el nivel de madurez de Definición, ni antes ni después de las mejoras. Teniendo en cuenta que para que un nivel pueda estar consolidado, es necesario que estén consolidados todos los niveles inferiores, es imposible que algún otro nivel superior pudiera estar consolidado para este PGI, con lo que no habría sido necesario seguir haciendo la evaluación. De todos modos se siguen exponiendo los resultados más significativos del resto de la evaluación.

Con respecto a las ACP del nivel 3, se obtuvieron las siguientes conclusiones:

- Verificación y Validación de los componentes del PGI y Productos de Información (VV). La organización realizó pruebas de verificación y validación para la aplicación SINCRO y para la base de datos que le daba soporte, pero no para ningún otro componente del sistema. Se calculó la calificación de "Parcialmente Satisfecho". Como medida de mejora se propuso desarrollar planes de verificación y validación para todos los componentes del PGI. Tras el tiempo de implementación de mejoras no se realizó ninguna de las actividades propuestas, con lo que la calificación continuó en "Parcialmente Satisfecho".
- Gestión del Impacto de los riesgos debido a una pobre calidad de datos y de Información (GIR). La organización no contemplaba ninguna gestión del riesgo ni antes ni después de la evaluación y propuestas de mejoras, con lo que en los dos momentos, este ACP obtuvo la calificación de "No Satisfecho".
- Gestión de la Estandarización de la Calidad de la Información (GECI). La organización contempla según su manual de calidad los procesos de estandarización de procesos y publicación de nuevos conocimientos adquiridos. Por tanto tiene la infraestructura adecuada para dar soporte a esta ACP. Lo único que se hace hasta el momento es transmitir en los cursos de formación a los nuevos empleados ciertas reglas de nomenclatura de datos y rutinas típicas de validación según el tipo de atributo usado. Se propuso como mejora utilizar esa infraestructura para la estandarización de la calidad de la información. Como los resultados obtenidos no fueron significativos, siguió con la misma calificación: "Parcialmente Satisfecho".
- Gestión de las Políticas Organizacionales de Calidad de la Información (GPCIO). Al igual que en la ACP anterior, la organización tiene una infraestructura adecuada para dar soporte a la gestión de Políticas Organizacionales sobre calidad de la información. Lo

que se ha venido haciendo al respecto ha sido tomar la decisión de decidir unificar criterios y nomenclaturas en el diseño de las bases de datos y que había que transmitir esa información a los nuevos trabajadores en los cursos de formación. Pero no se hace de forma sistemática. La calificación que se obtuvo fue la de "Parcialmente Satisfecho". Como mejora se propuso identificar y mejorar políticas de calidad de datos y de información que afectasen a todos los componentes del PGI. Tras el período de implementación de mejoras, la organización mantuvo la situación como estaba. Esto hace que la calificación de esta ACP continuase siendo "Parcialmente Satisfecho".

En lo que se refiere a las ACP de nivel 4 (Gestión Cuantitativa), se obtuvieron las siguientes conclusiones:

- Gestión de Planes de Medición (GPM). Como la organización no gestionaba para el PGI, ningún tipo de dimensión ni métrica de calidad de información, era imposible gestionar planes de medición. El estado de esta ACP es "No Satisfecho". Para mejorar esta situación se propuso elaborar planes de medida cuando estuviesen adecuadamente definidas las dimensiones y métricas de calidad de datos. No se observaron estas mejoras, por lo que esta ACP continua en el estado de "No Satisfecho".
- Gestión de la Automatización de los Planes de Medición (GAPM). Al igual que ocurría con la ACP anterior, es imposible satisfacer ésta, ya que ni siquiera estaban desarrollados los planes para la medición. El estado por tanto es el de "No Satisfecho". Como mejora se propuso automatizar los planes cuando los hubiese. Como no se observó esta posibilidad, fue imposible automatizar ningún plan. Como consecuencia el estado sigue siendo "No Satisfecho".
- Por último, y con respecto a las ACP del nivel 5 (Optimizante) el proceso de evaluación arrojó estos resultados:
- Análisis Causal para la prevención de defectos (GPD). Al no disponer de valores de métricas de calidad de datos y de información es imposible aplicar cualquier tipo de herramienta o técnica para identificar posibles causas de defectos. No obstante la organización cuenta con infraestructura y experiencia en el análisis de defectos debido al sistema de calidad. El estado de esta ACP es de "Parcialmente Satisfecho". Se recomendó que cuando se tuviesen medidas, aplicar técnicas y herramientas estadísticas para la identificación de las causas de los defectos. Desdichadamente, como tras el tiempo de implementación de mejoras no se tenían dimensiones de calidad ni métricas, ni planes para la medición, fue imposible realizar ningún tipo de análisis. El estado de esta ACP tras las propuestas de mejoras, seguía siendo "Parcialmente Satisfecho".
- Gestión de la Innovación y del Desarrollo Organizacional (GIDO). Esta ACP se basa en el estudio de los valores de las métricas para determinar qué aspectos del PGI se pueden mejorar. Dada la situación de no existencia de medidas de las

métricas, el resultado de la evaluación deja a esta ACP en el estado de "No Satisfecho". Como propuestas de mejoras se sugirió utilizar los valores de las métricas para descubrir qué aspectos del PGI podrían mejorarse para aumentar su rendimiento. Tras el período de mejoras, al no existir ni dimensiones de calidad, ni medidas, no se pudo satisfacer esta ACP.

5. CONCLUSIONES Y TRABAJO FUTURO.

En este artículo se ha presentado un marco de trabajo para la evaluación y mejora de la calidad de la información a través de la gestión. Consta de un Modelo de Calidad de Información basado en Niveles de Madurez (CALDEA) y una Metodología para la Evaluación y Mejora (EVAMECAL). Esta estructura satisface las cuatro condiciones mencionadas previamente: Como CALDEA está estructurado en niveles de madurez se proporciona un conjunto sistemático y conciso de criterios para evaluar la calidad de la información, satisfaciendo la primera condición. Definiendo ACPs (algunas orientadas a la gestión) para cada nivel se proporciona una base para la gestión proactiva, satisfaciendo así la tercera condición. Y finalmente, estando el modelo estructurado en niveles por fases (donde se agrupan ACPs, actividades, herramientas, técnicas y estándares) se facilita un mapa conceptual para que la comunidad investigadora pueda abordar una gran variedad de planteamientos, teorías y fenómenos relacionados con la calidad de la información, satisfaciendo la cuarta condición. Por otro lado, EVAMECAL proporciona un esquema para analizar y resolver problemas relacionados con la calidad de la información, satisfaciendo la segunda condición.

Para el futuro, tenemos varias intenciones: en primer lugar, finalizar nuestra línea actual de investigación, refinando CALDEA y EVAMECAL aplicándolas a diferentes clases de organizaciones (hemos aplicado ya el marco de trabajo a algunas organizaciones con más o menos éxito); por otro lado, queremos estudiar la utilidad de las técnicas, herramientas y prácticas actuales o emprender el desarrollo de las que no existan, para tener una colección de técnicas y herramientas recomendables para los distintos contextos.

6. AGRADECIMIENTOS.

Esta investigación es parte de los proyectos CALIPO financiado por la Dirección General de Investigación del Ministerio de Ciencia y Tecnología (TIC2003-07804-C05-03)- y MESSENGER financiado por la Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha (PCC-03-003-1).

7. REFERENCIAS.

- [1] Ballou D., Wang R., Pazer, H., y Tayi, G.K. Modeling Information Manufacturing Systems to Determine Information Product Quality. *Management Science* 44(4) (1998) Pp. 462-484,
- [2] Ballou, D. y Tayi, G.K. Enhancing data quality in Data Warehouse Environments. *Communications of the ACM*, Vol 42, No I. (1999)
- [3] Bicego, A., y Kuvaja, D. Bootstrap, Europe's Assesment Method, *IEEE Software*, (1993), Pp. 93-95.
- [4] Bobrowski, M., Marré, M., y Yankelevich, D. A Software Engineering View of Data Quality. *Proceedings of*

Second International Software Quality in Europe. (Bélgica, 1998)

[5] Bouzeghoub, M. y Kedad, Z. A quality-based framework for Physical Data Warehouse Design". Proceedings of the International Workshop on Design and Management of Data Warehouse (DMDW'2000) (Estocolmo, Suecia, Junio 5-6, 2000)

[6] Bringel, H., Caetano, A., Tribolet, J. Business Process Modeling Towards Data Quality Assurance. Proceeding of ICEIS'2004. (Oporto, Portugal 2004) Pp 565-568

[7] Burgess, M.S.E., Gray, W.A., Fiddian, N.J. A flexible quality framework for use within information retrieval Proceedings of the Eighth International Conference on Information Quality (ICIQ'2003). Pp 297-313. 2003

[8] Calero, C. y Piattini, M. Metrics for databases: a way to assure the quality. In: Information and Database Quality. Kluwer Academic Publishers. Pp. 57-84. 2002

[9] Coallier, F. How ISO 9001 fits into the software world, IEEE Software. (Enero 1994) Pp 98 -100.

[10] Crosby, P.B. Quality is free. (Nueva York, Nueva York: McGraw Hill, 1979)

[11] Deming, W.E. Out of the crisis. (Cambridge, MA. MIT Center for advanced engineering study, 1986)

[12] Dunaway, D. K. CMM SM -Based Appraisal for Internal Process Improvement. (CBA IPI) Lead Assessor's Guide (CMU/SEI-96-HB-003). (Software Engineering Institute, Carnegie Mellon University, Pittsburgh, 1996)

[13] English, L.P. Improving Data Warehouse and Business Information Quality: Methods for reducing costs and increasing Profits. Willey & Sons, (1999).

[14] English, L.P. Total Quality data Management (TQdM), Methodology for Information Quality Improvement. In: Information and Database Quality. Kluwer Academic Publishers. (2002).Pp. 85-110.

[15] Eppler, M. J. Managing Information Quality. Springer. (Alemania, 2003).

[16] Eppler, M. J. and Wittig, D. Conceptualizing Information Quality: A review of Information Quality Frameworks from the last ten years." Proceedings of the 2000 Conference on Information Quality. Pp 83-96

[17] Fagan, M. Design and Code Inspections to Reduce Errors in Program Development. IBM Systems Journal 15, 3 Pp 182-211. 1976

[18] Florac, W, A.and Carleton, A.D. Using Statistical Process Control to Measure Software Process. In Fundamental Concepts for the Software Quality Engineer. Taz Daughtrey Editor. American Society for Quality. Pp.133-144. 2002.

[19] Fuggeta, A. Software Process: A roadmap. The future of Software Engineering, ed. A. Finkelstein ACM, Press. (2000) Pp.27-34.

[20] Gertz, M., Tamer, M., Saake, G. Sattler, K., Report on the Dagstuhl Seminar "Data Quality on the web". Proceedings of SIGMOD record. Vol.33, No.1. (Marzo 2004)

[21] Getto G. (2002) Risk Management Supporting Quality Management of Software Acquisition Projects. In Fundamental Concepts for the Software Quality Engineer. Taz Daughtrey Editor. American Society for Quality. (2002) Pp. 25-38

[22] Gilb, T. and Graham D. Software Inspection. London: Addison-Wesley Longmain. (1993)

[23] Grimmer, U., and Hinrichs, H. A methodological approach to data quality management supported by data

mining. Proceedings of the Sixth International conference on Information Quality.

(2001) Pp 217-232.

[24] Hinrichs, H. CLIQ- Intelligent Data Quality Management. Proceedings of the fourth IEEE international Baltic Workshop on databases and Information System. (2000)

[25] Hinrichs, H. and Aden, T. An ISO 9001:2000 compliant quality management System for Data Integration in Data Warehouse System. Proceedings of the International Workshop on Design and Management of Data Warehouse (DMDW'2001) (Interlaken, Suiza, Junio 4, 2001)

[26] Hoxmaier, J.A. Dimensions of Database Quality. In Developing Quality Complex Database Systems: Practices, Techniques, and Technologies. Editor Shirley Becker. Idea Group Publishing. (2001)

[27] Huang, K.T., Lee, Y., Wang, R. Quality Information and Knowledge. Prentice Hall, Upper Saddle River. (1999)

[28] Humphrey, W. Managing the software process, Addison Wesley, Reading Mass. (1989)

[29] Humphrey, W. The Software Standard Profile. In Fundamental Concepts for the Software Quality Engineer. Taz Daughtrey Editor. American Society for Quality. (2002) Pp. 3-16.

[30] IEEE STD 1012-1986 IEEE Standard for Software Verification and Validation Plans.

[31] IEEE STD 1058.1-1987 IEEE Standard for Software Project Management Plans.

[32] IEEE STD 1061-1992. IEEE Standard for a Software Quality Metrics Methodology.

[33] IEEE STD 830-1998. IEEE Standard guide to Software Requirements Specification. (IEEE Nueva York, EEUU IEEE Computer Society).

[34] ISO/IEC 9126: Information technology Software product evaluation Quality characteristics and guidelines for their use, ISO/IEC Standard 9126.(Genova, 2001).

[35] ISO IEC 15504 TR2: 1998, part 2: A reference model for processes and process capability, ISO/IEC JTC1/SC7, 1998.

[36] ISO IEC 15504 TR2: 1998, Software Process Assessment Parte 7: Guide for use in process improvement, ISO/IEC JTC1/SC7.(1998).

[37] Jarke, M., y Vassiliou, Y. (1997). Data Warehouse Quality: A review of the DWQ Project. Proceedings of Second Conference on Information Quality. (Massachusetts Institute of Technology, Cambridge)

[38] Juran, J.M. Juran on Planning for Quality. Nueva York: McMillan. (Nueva York, 1988)

[39] Kahn, B., Strong, D., Wang, R. "Information Quality Benchmarks: Product and Service Performance". Communications of the ACM (Abril 2002/Vol. 45, No. 4)

[40] Kim, W., Choi, B., Towards Quantifying Data Quality Costs in Journal of Object Technology, Vol. 2, no.4.(Julio-Agosto 2003) Pp. 69-76.

[41] Kiszurno, E., Yankelevich, D., Data Testing. ASSE 2001 Proceedings of SADIO. (2001)

[42] Lee, Y.W., Strong, D., Kahn, B., Wang, R. AIMQ: a methodology for information quality assessment. Information & Management. (2001)

[43] Loshin D. Enterprises Knowledge Management: The Data Quality Approach. Morgan Kauffman. (San Francisco, California, 2001)

[44] Meredith, D.C. Managing with Metrics: Theory into

Practice. In *Fundamental Concepts for the Software Quality Engineer*. Taz Daughtrey Editor. American Society for Quality. (2002) Pp145-154

[45] Olson, J. E. *Data Quality: the accuracy dimension*. Ed. Morgan Kaufmann Publishers. (2003)

[46] Paulk, M., C. Weber, B. Curtis, and Chrissis, M. *The Capability Maturity Model Guideline for Improving the software Process*, Addison-Wesley, Reading, Mass. (1995)

[47] Pipino, L., Lee, Y., and Wang, R. *Data Quality Assessment*. *Communications of the ACM*. (Abril 2002/Vol. 45, No. 4)

[48] Redman, T.C. *Data Quality for the Information Age*. Artech House Publishers. (Boston, 1996)

[49] Redman, T.C. *Data Quality: The Field Guide*. Digital Press. (2001)

[50] SEI. *Capability Maturity Model® Integration (CMMISM), Version 1.1 CMMISM (CMMI-SE/SW/IPPD/SS, V1.1) Staged Representation CMU/SEI-2002-TR-012 ESC-TR-2002-012*. In <http://www.sei.cmu.edu/publications/documents/02.reports/02tr002.html> (ultimo actualización en Junio de 2004)

[51] Shankaranarayanan, G. Wang, R., Ziad, M. *IP-MAP: Representing the Manufacture of an Information Product*. *Proceedings of the 2000 Conference on Information Quality*. (2000)

[52] Smith, C., and Heights, A. *Defect Prevention: The road less Traveled*. In *Fundamental Concepts for the Software Quality Engineer*. Taz Daughtrey Editor. American Society for Quality. (2002)

[53] *Standard CMMI SM Appraisal Method for Process Improvement (SCAMPI SM), Version 1.1: Method Definition Document*. CMU/SEI-2001-HB-001.

[54] Strong, D.M., Lee, Y. W., Wang R.Y. *Data Quality in context*. *Communications of the ACM*. (Mayo 1997) pp 103-110.

[55] Van Soligen, R. and Berghout, E. *The Goal/Question/Metric Metodology: a practical guide for quality improvement of software development*. Ed McGraw-Hill. (1999)

[56] Wang, R., *A product perspective on data quality management*. *Communications of the ACM*. Vol 41(2) (Febrero 1998) pp58-65.

[57] Wang, R., Kon, H., and Madnick S. *Data Quality Requirements Analisis and Modeling*. Published in the Ninth International Conference of Data Engineering Vienna, Austria. (1993) Pp 670 677.

[58] Wang, R., Strong, D., Guarascio, L.M. *Beyond Accuracy: What Data Quality Means to Data Consumers*. *Journal of Management Information Systems*, , Volume 12, No. 4. (Primavera 1996) pp. 5-33.

[59] Wang, R., Ziad, M., Lee, Y.W. *Data quality*. Kluwer Academic Publishers. (Massachussets (EEUU), 2001)

