

**UNIVERSIDAD CATÓLICA DE LA SANTÍSIMA CONCEPCIÓN
FACULTAD DE INGENIERÍA
INGENIERÍA CIVIL INFORMATICA**



**“PLAN DE DATA GOVERNANCE PARA LA
DIRECCIÓN DE ADMISIÓN Y REGISTRO
ACADÉMICO DE LA UNIVERSIDAD CATÓLICA DE
LA SANTÍSIMA CONCEPCIÓN”**

Camila Fernanda Coloma Boettcher

Informe de Proyecto de Título para optar al Título de Ingeniero Civil
Informático

Profesor Guía

Dra. Emma Chávez.

Concepción, Agosto del 2017

Resumen

Hoy en día los datos son un activo valioso para las organizaciones, ya que estos con la ayuda de los sistemas de información (SI) proveen de información útil para la toma de decisiones, siendo un importante insumo para el plan de negocios y la planificación estratégica.

Un uso adecuado de los datos genera ventajas competitivas y marca diferencias en el mercado, la cantidad de datos capturados alcanza límites que el ser humano no es capaz de procesar de forma manual, por lo que asegurar la seguridad, calidad y claridad de los datos es imprescindible.

Con el fin de mantener un uso adecuado y estratégico de los datos, Data Governance (DG), controla los procedimientos asociados a los datos, para que así la organización pueda hacer un uso eficaz de la información logrando así sus objetivos y ventajas competitivas (Fragoso, 2013). DG ayuda a mejorar la toma de decisiones, proteger los datos, apoya al personal en la adopción de enfoques comunes, crea y mejora estándares, reduce costos y aumenta la eficacia de los datos. DG no solo se encarga de los datos, sino que también incluye personas, procesos y tecnologías; puede centrarse en distintas áreas de la organización tales como políticas, calidad de datos, seguridad, arquitectura, inteligencia de negocios y apoyo a la gestión.

En este proyecto se analizaron el sistema, los datos, responsables y la documentación del sistema académico curricular SIMBAD que administra la Dirección de Admisión y Registro Académico (DARA) de la Universidad Católica de la Santísima Concepción. Se generó por tanto un plan de DG enfocado al área de calidad y seguridad de los datos. Se establecieron un conjunto de procesos y políticas para la obtención y utilización de los datos; la clarificación de roles y responsabilidades; y se definieron una serie de dimensiones con las cuales es posible realizar un control de la calidad y seguridad de los datos a lo largo del tiempo.

Abstract

Today data is a valuable assets for organizations, given that information systems (IS) provide of useful information for decision making, being an important input to the business and strategic planning.

A suitable use of data generates competitive advantages and marks differences on the market, the quantity of captured data reaches limits that the human being is not capable of processing in a manual form. Therefore to assure the security, quality and clarity of data is essential.

In order to maintain a suitable and strategic use of data, Data Governance (DG) controls the procedures associated with data, so that the organization could make an effective use of the information by achieving through this way its targets and competitive advantages (Fagoso, 2013). DG helps to improve the decision making, protects data, supports to the personnel in the adoption of common approaches, creates and improves standards, reduces costs and the efficacy of data increases. DG not only is responsible for data, but also it includes individuals, processes and technologies; it focuses on different areas of the organization such as politics, data quality, security, architecture, business intelligence and support management.

In this project there were analyzed the system, data, responsible users and the documentation of the academic curricular system SIMBAD who administers the Dirección de Admisión y Registro Académico (DARA) of the Universidad Católica de la Santísima Concepcion. Therefore it was developed a plan of DG focused on data quality and safety of data. There was established a set of processes and politics for the safety and use of the data; the clarification of roles and responsibilities; and there were defined a number of dimensions to achieve a control of the quality and safety of the data.

Dedicatoria

Dedico este proyecto de título a mis padres, Arturo y Ninette, por el gran esfuerzo, amor, consejos y apoyo que me han otorgaron durante toda mi carrera estudiantil; a mi pololo Cristopher que me ha acompañado en todo este periodo, entregándome todo su amor y un apoyo incondicional en todo momento; a toda mi familia y amigos por haber creído en mí.

Quiero agradecer a cada uno de los profesores con que tuve el agrado de compartir, ya que la entrega de sus conocimientos ha sido de vital ayuda para mi formación como profesional y en especial a mi profesora guía Emma por su paciencia y dedicación.

Índice General

Capítulo 1 . Introducción.....	1
1.1 Presentación del tema.....	1
1.2 Objetivos	1
1.2.1 Objetivo general	1
1.2.2 Objetivos específicos	1
1.3 Justificación del problema.....	2
1.4 Delimitación del problema	2
1.5 Metodología.....	3
Capítulo 2 . Marco teórico.....	4
2.1 Dato	4
2.2 Información	4
2.3 Conocimiento	5
2.4 Calidad de datos	6
2.5 Seguridad de los datos	8
2.6 Gestión de datos	9
2.7 Gestión de la calidad de los datos	9
2.8 Sistemas de información	10
2.9 Data Governance (DG).....	11
2.10 Dirección de Servicios Informáticos	13
2.11 Sistema académico curricular (SIMBAD).....	13
Capítulo 3 . Estado del arte	15
3.1 Gestión de la seguridad de la información	15
3.1.1 Principios para una correcta seguridad de la información	15
3.1.2 Un marco para la seguridad de la información	16
3.1.3 Un modelo dinámico para la seguridad de la información	17
3.2 Gestión de Data Governance	17
3.2.1 Un modelo para Data Governance	18
3.2.2 Data Governance con enfoque ágil	19
3.2.3 Una morfología para Data Governance.....	20
3.2.4 Áreas de aplicación de Data Governance.....	22

3.2.5	Dificultades al aplicar Data Governance.....	24
3.2.6	¿Por qué usar Data Governance?, principales resultados.....	24
3.3	Conclusión y discusión del capítulo.....	25
	Capítulo 4 . Diagnóstico de calidad y seguridad de los datos	30
4.1	Entrevistas	30
4.2	Encuesta de calidad de los datos	31
4.3	Revisión del sistema en cuanto a calidad de los datos	38
4.4	Encuesta de seguridad de los datos	43
4.5	Revisión del sistema en cuanto a seguridad de los datos	47
4.6	Conclusión y discusión del capítulo.....	49
	Capítulo 5 . Procedimientos para la calidad y seguridad de los datos.....	51
5.1	Procedimientos de calidad de los datos	51
5.1.1	Compleitud.....	51
5.1.2	Consistencia	56
5.1.3	Exactitud	58
5.1.4	Unicidad	63
5.1.5	Validez	65
5.2	Procedimientos de seguridad de los datos	68
5.2.1	Confidencialidad	68
5.2.2	Integridad	69
5.2.3	Disponibilidad.....	70
5.3	Conclusión y discusión del capítulo.....	71
	Capítulo 6 . Políticas para el sistema SIMBAD.....	72
6.1	Políticas de acceso.....	72
6.1.1	Roles del sistema.....	72
6.1.2	Solicitudes de permisos en el sistema	74
6.2	Políticas de documentación	74
6.3	Políticas de mantención.....	75
6.4	Políticas de respaldo.....	76
6.5	Conclusión y discusión del capítulo.....	77
	Capítulo 7 Conclusión y recomendaciones	78
	Bibliografía	81

Anexos	85
Anexo A. Catálogos de sistemas.....	85
Anexo B. Encuesta.....	90
Anexo C. Gráficos de las encuestas realizadas.....	93
Anexo D. Datos activos para el sistema SIMBAD	99

Índice de Figuras

Figura 2-1: Esquema información (Fuente: Propia).....	4
Figura 2-2: Dato, información, conocimiento (Fuente: Propia).....	6
Figura 2-3: Dimensiones para evaluar calidad de los datos (UK, DAMA, 2013)).....	7
Figura 2-4: Proceso de los sistemas de información (Fuente: Propia).....	10
Figura 2-5: Data Governance (Fuente: Propia).....	12
Figura 3-1: Data Governance ágil (Fuente: Propia).....	20
Figura 3-2: Una morfología para la organización de Data Governance (Otto, 2011).....	21
Figura 5-1: Proceso de asignación de correo institucional (Fuente: Propia)	53
Figura 5-2: Ejemplificación de la estandarización de los datos (Fuente: Propia).....	60
Figura 5-3: Ejemplificación de la descripción de los datos en el campo sexo (Fuente: Propia).....	63

Índice de Tablas

Tabla 2-1: Características de la información (Fuente: Propia).	5
Tabla 2-2: Principales módulos de SIMBAD (Fuente: Propia).	14
Tabla 3-1: Proyecto de un modelo de Data Governance (Wende, A Model for Data Governance – Organising Accountabilities for data quality management, 2007).	19
Tabla 3-2: Comparación de los enfoques de la gestión de seguridad de la información (Fuente: Propia).....	27
Tabla 3-3: Comparación de los enfoques Data Governance (Fuente: Propia).....	28
Tabla 3-4: Resumen de implementaciones (Fuente: Propia).	29
Tabla 4-1: Revisión a la base de datos de SIMBAD (Fuente: Propia).....	40
Tabla 4-2: Tabla PERSONAS (Dirección de Servicios Informaticos, 2007).	42
Tabla 4-3: Revisión a la base de datos de SIMBAD (Fuente: Propia).....	48
Tabla 6-1: Rol de mantención del sistema SIMBAD (Fuente: Propia).	72
Tabla 6-2: Roles de uso del sistema SIMBAD (Fuente: Propia).	73
Tabla A-1: Sistemas del área académico curricular (Dirección de Servicios Informáticos, 2016).	85
Tabla A-2: Sistemas del área administrativo (Dirección de Servicios Informáticos, 2016).	88
Tabla D-3: Tablas principales para el sistema (Fuente: Propia).	99

Índice de Gráficos

Gráfico 4-1: Resultado de la encuesta en completitud de los datos (Fuente: Propia).....	34
Gráfico 4-2: Resultado de la encuesta en consistencia de los datos (Fuente: Propia).	35
Gráfico 4-3: Resultado de la encuesta en exactitud de los datos (Fuente: Propia).	35
Gráfico 4-4: Resultado de la encuesta en unicidad de los datos (Fuente: Propia).	36
Gráfico 4-5: Resultado de la encuesta en oportunidad de los datos (Fuente: Propia).....	37
Gráfico 4-6: Resultado de la encuesta en validez de los datos (Fuente: propia).....	37
Gráfico 4-7: Resultados encuesta calidad de los datos (Fuente: propia).	38
Gráfico 4-8: Resultados revisión base de datos de SIMBAD (Fuente: Propia).	42
Gráfico 4-9: Resultado en cuanto a confidencialidad de los datos (Fuente: Propia).	45
Gráfico 4-10: Resultado en cuanto a integridad de los datos (Fuente: Propia).....	46
Gráfico 4-11: Resultado en cuanto a disponibilidad de los datos (Fuente: Propia)	46
Gráfico 4-12: Resultado encuesta seguridad de los datos (Fuente: propia).	47
Gráfico 4-13: Resultados revisión base de datos de SIMBAD (Fuente: Propia).....	49
Gráfico C-0-1: Funciones realizadas por los usuarios en el sistema (Fuente: Propia).....	93
Gráfico C-0-2: Resultados por pregunta en completitud de los datos (Fuente: Propia).	94
Gráfico C-0-3: Resultados por pregunta en consistencia de los datos (Fuente: Propia).	94
Gráfico C-0-4: Resultados en cuanto a exactitud de los datos (Fuente: Propia).....	95
Gráfico C-0-5: Resultados por pregunta en unicidad de los datos (Fuente: Propia).....	95
Gráfico C-0-6: Resultados por pregunta en oportunidad de los datos (Fuente: Propia).	96
Gráfico C-0-7: Resultados por pregunta en validez de los datos (Fuente: Propia).	97
Gráfico C-0-8: Resultados por pregunta en confidencialidad de los datos (Fuente: Propia).	97
Gráfico C-0-9: Resultados por pregunta en integridad de los datos (Fuente: Propia).	98
Gráfico C-0-10: Resultados por pregunta en disponibilidad de los datos (Fuente: Propia).	99

Abreviaciones

Abreviaciones	
SIMBAD	Sistema Académico Curricular.
UCSC	Universidad Católica de la Santísima Concepción.
DARA	Dirección de Admisión y Registro Académico
DYSSAE	Unidad de Desarrollo y Soporte de Sistemas Académicos Estudiantiles.
DG	Data Governance.
CUBIX	Plataforma Inteligencia de Negocio.
SI	Sistemas de Información
TI	Tecnologías de Información

Capítulo 1 . Introducción

1.1 Presentación del tema

En la actualidad los datos son parte fundamental en las organizaciones, ya que estos son de gran apoyo para la toma de decisiones. Es por esto que el mantener un uso adecuado de los datos es de suma importancia, la realización de un plan de Data Governance (DG), el cual busca asegurar el uso eficaz y eficiente de los datos con el fin de que las organizaciones logren alcanzar sus objetivos. Un plan de DG involucra no solo los datos, sino que también a las personas, los procesos y las tecnologías, estos planes pueden centrarse en distintas áreas: políticas, calidad de datos, seguridad, arquitectura, inteligencia de negocios, apoyo a la gestión, entre otros (Gwen, 2015).

En este proyecto se analizó el sistema, los datos, responsables y la documentación del sistema SIMBAD administrado por la DARA, sistema responsable de registrar todos los procesos académicos de los alumnos de la Universidad Católica de la Santísima Concepción. Se generó por tanto un plan de DG enfocado al área de calidad y seguridad de los datos de este sistema, con lo cual se estableció un conjunto de procesos y políticas que apoyan el cumplimiento de una serie de dimensiones propuestas, además se esclarecieron roles y responsabilidades que van en ayuda del mejoramiento y correcto funcionamiento del sistema.

1.2 Objetivos

1.2.1 Objetivo general

Diseñar un plan de Data Governance basado en aspectos de seguridad y calidad de los datos del sistema SIMBAD de la Dirección de Admisión y Registro académico (DARA), de la Universidad Católica de la Santísima Concepción.

1.2.2 Objetivos específicos

1. Estudiar planes de Data Governance ya realizados y la información relativa al área de estudio.

2. Diagnosticar estado actual del sistema SIMBAD de la DARA
3. Definir reglas, objetivos de calidad y definición de datos para el sistema SIMBAD.
4. Definir políticas y responsabilidades para el sistema SIMBAD.
5. Documentar el trabajo realizado.

1.3 Justificación del problema

La Dirección de Admisión y Registro Académico actúa como una garantía para el alumnado, ya que resguarda y administra los antecedentes personales y académicos de los estudiantes de la UCSC (UCSC, 2016). La DSI de la UCSC pone a disposición de la DARA tecnologías de información como el sistema SIMBAD que va en ayuda de las necesidades de la comunidad universitaria (Dirección de Servicios Informáticos de la UCSC, 2016).

Se vislumbra, de acuerdo a información entregada por usuarios, que parecieran existir inconsistencias de datos en el sistema SIMBAD, lo que puede deberse a un problema de duplicidad de datos, errores de ingreso, que no exista una normativa que rija el uso o carga de datos o que no se defina quien es el responsable de estos, entre otros.

Desarrollar un plan de DG enfocado a la calidad y seguridad de los datos del sistema SIMBAD ayuda a que este cumpla con las garantías que propone la DARA, además de resguardar y administrar los antecedentes personales y académicos como también de clarificar el proceso y políticas del uso y obtención de los datos para dicho sistema.

1.4 Delimitación del problema

El proyecto se delimita por:

1. Analizar la calidad y seguridad de los datos del sistema SIMBAD.
2. El proyecto solo analiza políticas y aspectos de seguridad y calidad de datos del sistema SIMBAD de la Universidad Católica de la Santísima Concepción.

1.5 Metodología

- Por medio de una revisión bibliográfica, se revisaron planes de Data Governance ya realizados e información de interés para el proyecto, esta se documentó junto a los conceptos esenciales a comprender en el área de estudio. Los artículos revisados fueron extraídos de fuentes tales como: Google Scholar y ScienceDirect.
- A través de entrevistas realizadas a los encargados de la Dirección de Servicios Informáticos, se realizó un diagnóstico actual de los sistemas en cuanto a calidad y seguridad de los datos.
- Realizando encuestas, revisión de la documentación del sistema y una revisión y análisis de las políticas, estándares y definición de los datos existentes, se procedió a identificar las reglas existentes, eliminar contradicciones, dar prioridad a algunas normas y según fuera necesario establecer algunas reglas.
- Luego de la investigación realizada previamente, se definieron los encargados de la toma de decisiones en relación a los datos del sistema identificando los responsables de cada actividad realizada con estos.
- Todo lo realizado anteriormente fue documentado.

Capítulo 2 . Marco teórico

Para poder abordar el tema en estudio es necesario aclarar algunos conceptos, los cuales se presentan en este capítulo.

2.1 Dato

Los datos son la fuente primaria de la información, son elementos discretos, hechos que no presentan juicio o interpretaciones que por sí mismos no tienen un valor (Giner, 2004).

2.2 Información

La información es un grupo de datos procesados, con un orden lógico y contextualizado, los cuales describen sucesos o entidades, en la Figura 2-1 es posible apreciar un esquema de lo anteriormente descrito (Instituto Superior Nuestra Sra. de la Paz, 1016).

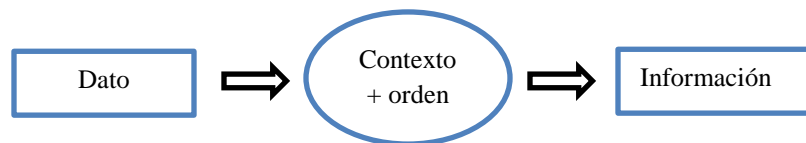


Figura 2-1: Esquema información (Fuente: Propia).

Según Idalberto Chiavenato (Chiavenato, 2006), información "Es un conjunto de datos con un significado, que reduce la incertidumbre o que aumenta el conocimiento de algo. En verdad, la información es un mensaje con significado en un determinado contexto, disponible para uso inmediato y que proporciona orientación a las acciones por el hecho de reducir el margen de incertidumbre con respecto a nuestras decisiones".

Es completamente distinto referirse a datos o a información ya que, un conjunto del primero no significa nada por sí solo, pero si estos son procesados y contextualizados se convierten en información (Alvin, 2006). Para que la información sea de utilidad en la toma de decisiones, esta debe tener las características que se muestran en la Tabla 2-1, las cuales le otorgan mayor utilidad (Stair & Reynolds, 2010).

Tabla 2-1: Características de la información (Fuente: Propia).

CARACTERÍSTICAS DE LA INFORMACIÓN	
Características	Definiciones
Accesible	Debe poder ser accesible de una manera intuitiva, de tal forma que puedan obtenerla en el formato correcto y en el tiempo destinado para satisfacer las necesidades requeridas.
Exacta	Cuando es exacta, la información está libre de errores. En algunos casos se genera información imprecisa debido a que el proceso de transformación es alimentado con datos erróneos o no pertinentes.
Completa	La información completa contiene todos los hechos relevantes.
Flexible	La información es flexible cuando puede utilizarse para una gran variedad de propósitos.
Relevante	Es relevante cuando es importante para la toma de decisiones.
Confiable	Pues depende de la veracidad de su fuente de información.
Verificable	La información debe ser verificable. Se podrá comprobar con el fin de asegurar de que es correcta.

Las fuentes de información pueden ser:

- Fuente de información primaria: Información que ha sido publicada por primera vez, que no ha sido interpretada o evaluada por nadie más.
- Fuente de información secundaria: Información primaria sintetizada y reorganizada que facilita el acceso a los contenidos.
- Fuente de información terciaria: Guías físicas o virtuales que contienen información sobre las fuentes secundarias.

2.3 Conocimiento

Para Davenport y Prusak (Negocio, 2016) el conocimiento es: “Una mezcla de experiencia, valores, información y “saber hacer” que sirve como marco para la

incorporación de nuevas experiencias e información, y es útil para la acción. Se origina y aplica en la mente de los concedores. En las organizaciones con frecuencia no sólo se encuentra dentro de documentos o almacenes de datos, sino que también está en rutinas organizativas, procesos, prácticas, y normas”.

El conocimiento es derivado de la información, donde el primero solo puede ser obtenido de un interpretador o concedor del tema, donde este realiza comparaciones, predicciones y conexiones de la información generando así conocimiento (Prada, 2008). En la Figura 2-2 es posible apreciar la diferencia entre dato, información y conocimiento.

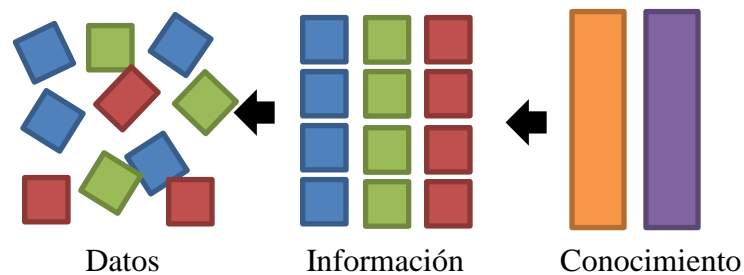


Figura 2-2: Dato, información y conocimiento (Fuente: Propia).

2.4 Calidad de datos

Puede definirse como el grado de excelencia que presenta un dato en relación a su representación (Lucero Ascencio, 2016). Hablamos de calidad si el dato se adecúa a su uso esperado, en cuanto a operaciones, toma de decisiones y planificación en cuanto a las reglas definidas para estos (Juran, 1998). Las reglas no son más que condiciones que los datos deben cumplir para ser de calidad. Según Lucero (Lucero Ascencio, 2016) existen 6 dimensiones de calidad, como se muestra en la Figura 2-3:

- **Completitud:** Se refiere a que los datos deben estar íntegros. A esta dimensión se le relaciona el atributo de validez, ya que si un campo es obligatorio su porcentaje de completitud será un 100%, pero no necesariamente el dato ingresado debiese ser correcto. Por lo que completitud no asegura correctitud.

- **Consistencia:** Se refiere a la ausencia de diferencias entre comparaciones de representaciones distintas con los mismos datos.
- **Exactitud:** Grado en que los datos describen correctamente lo solicitado, también abarca el grado de estandarización de los datos.
- **Unicidad:** Se refiere a que los datos no deben estar duplicados, deben ser obtenidos de una sola fuente y por consiguiente deben ser consistentes.
- **Oportunidad:** Es el grado en el que los datos representan la realidad necesaria en un punto de tiempo, tiene relación con la exactitud.
- **Validez:** Se refiere a si los datos se ajustan a la sintaxis definida con anterioridad, ya sea formato o tipo de dato.

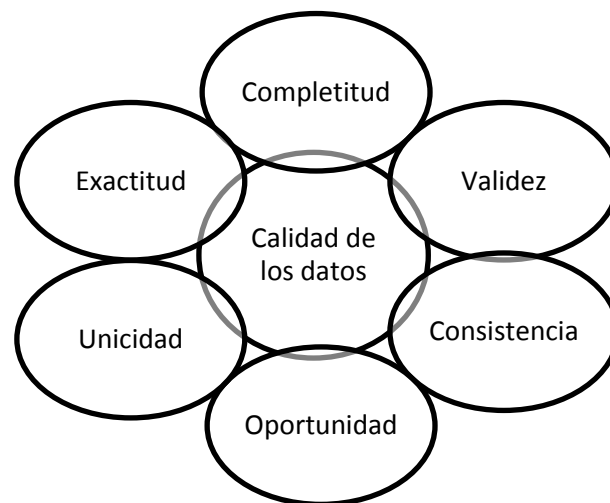


Figura 2-3: Dimensiones para evaluar calidad de los datos (DAMA, 2013)).

Las dimensiones mencionadas anteriormente son muy similares a las utilizadas para medir la calidad de información. Sin embargo, los datos son la materia prima de la información por lo que también estas dimensiones son utilizadas por diversos autores para evaluar la calidad de los datos en las organizaciones, incluyendo o quitando alguna de las propiedades con el fin de adecuarse a las ponderaciones asociadas a las distintas organizaciones (Evora, 2009), (DAMA, 2013). Calidad de datos es un concepto muy amplio, el cual no puede ser definido a menos que se establezcan los requisitos que se evaluarán en

ellos. Para este proyecto, se definirá como el dato correcto en el lugar apropiado, adquirido por la fuente adecuada, realizando la función establecida, única, consistente y completa. Si nos referimos a un proyecto en una organización, calidad de datos debe aplicarse en todas las fases, si una organización desea mejorar la calidad de datos, tiene que asegurar que existe un uso para cada uno de estos (Needham & Sinopoli, 2009) (Covella, 2005).

Para la evaluación de calidad de datos en el presente proyecto, se utilizó como base las 6 dimensiones propuestas anteriormente, estas se consideraron como las más apropiadas para ser evaluadas en la generación de un plan de DG.

2.5 Seguridad de los datos

La seguridad de los datos se refiere a la protección de los datos contra el acceso no autorizado, del uso, el cambio, la divulgación y la destrucción (Carnegie Mellon University , 2016). Muchas personas, grandes y pequeñas empresas dependen en gran medida de sistemas informáticos, que los datos de estos sean dañados, perdidos o robados, les generaría grandes pérdidas (BBC, 2016).

La ISO27001 (Gillies, 2011), estándar para la seguridad de la información plantea que las dimensiones a medir para evaluar la seguridad de los datos son: la confidencialidad, integridad y disponibilidad de los datos importantes para las organizaciones, a continuación se detallan cada uno de estos parámetros.

- **Confidencialidad:** Es la propiedad de prevenir la divulgación de información a personas o sistemas no autorizados.
- **Integridad:** La integridad es el atributo que busca mantener a los datos libres de modificaciones no autorizadas.
- **Disponibilidad:** Es la característica, cualidad o condición de los datos de encontrarse a disposición de quienes deben acceder a ella, ya sean personas, procesos o aplicaciones.

Luego de la revisión bibliográfica realizada se puede definir qué gran cantidad de autores basan sus estudios de seguridad de información y seguridad de los datos en la

ISO27001 (Sheikhpour & Modiri, 2012), (Susanto & Nabil, 2012); ya que consideran que los atributos propuestos abarcan todos los puntos de estudio necesarios para poder dar un correcto diagnóstico en cuanto a seguridad.

En el presente proyecto se utilizó como criterio de evaluación de la seguridad de los datos del sistema SIMBAD, las normas descritas en la ISO27001.

2.6 Gestión de datos

DAMA International (Copro, 2016), define gestión de datos como: "Gestión de recursos de datos es el desarrollo y ejecución de arquitecturas, políticas, prácticas y procedimientos que gestionan adecuadamente las necesidades del ciclo de vida de todos los datos de una empresa".

La gestión de datos se relaciona con el ciclo de vida de los datos, desde su creación hasta su eliminación, minimizando riesgos y costos de incumplimiento de normas, proporcionando acceso a los datos precisos en el momento necesario, sin generar ambigüedades o conflicto alguno, generando una buena comunicación. Se puede aplicar a cualquier dato de la organización que cuente con principios y procedimientos de la gestión de datos, es muy útil para fuentes de datos duplicados, ya que evita ambigüedades y asegura que los datos sean usados de la mejor manera en cuanto a acceso, almacenamiento, copias de seguridad y la jubilación, entre otras cosas (Informática, 2016).

Los beneficios que tiene la gestión de datos es mejorar el uso de los datos, entregando mayor seguridad a las organizaciones.

2.7 Gestión de la calidad de los datos

La gestión de la calidad de los datos se centra en recopilar, organizar, almacenar, procesar y presentar datos de alta calidad (Wende, A Model for Data Governance – Organising Accountabilities for Data, 2007). En otras palabras la gestión de calidad de datos se preocupa de la creación y despliegue de las funciones, las responsabilidades, las políticas y los procedimientos de adquisición, conservación, difusión y disposición de los datos (Geiger, 2004).

Para que una gestión de datos tenga éxito es esencial una asociación entre los grupos empresariales y la tecnología, ya que el primero es responsable de restablecer las reglas del negocio que regirán en los datos y de verificar la calidad de estos, en cambio la tecnología de información es responsable de establecer el medio ambiente en general, la construcción y el control de todo el entorno, es decir, la arquitectura , los sistemas, los establecimientos técnicos y bases de datos (Geiger, 2004).

Actualmente las organizaciones que se inician en la gestión de calidad de los datos, deben incluir descripción de los datos, limpieza, gestión de datos maestros, y principalmente DG como parte de su estrategia para así poder llevar a cabo una exitosa gestión de calidad de los datos. La gestión de calidad datos se ha vuelto cada vez más importante ya que el volumen de los datos se está expandiendo cada vez más rápido y las organizaciones buscan explotar su información para poder obtener una ventaja competitiva y mantener un control total sobre la integridad de la información en la empresa (Information Builders, 2016).

2.8 Sistemas de información

Un sistema de información (SI) es un conjunto de elementos o componentes interrelacionados que reúnen (entrada), manipulan (proceso), almacenan y distribuyen (salida) datos e información y proporciona una reacción correctiva (mecanismo de retroalimentación) si no se ha logrado cumplir un objetivo previamente definido.

El mecanismo de retroalimentación es el componente que ayuda a las organizaciones a cumplir sus objetivos (Stair & Reynolds, 2010) (ver Figura 2-4).

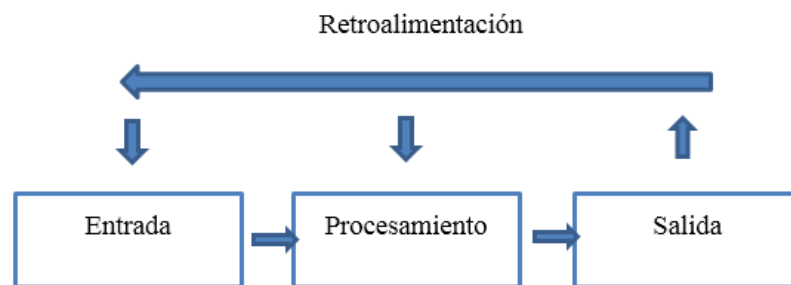


Figura 2-4: Proceso de los sistemas de información (Fuente: Propia).

Los sistemas de información fueron creados para alcanzar uno o más objetivos, ya que son un conjunto de elementos interrelacionados que buscan alcanzar un propósito u objetivo, también se pueden definir como una serie de procedimientos de datos, que escoge, almacena y recupera datos para disminuir la incertidumbre en la toma de decisiones mediante suministro de información a los niveles de la organización, inferior, medio y superior (Cauca, 2016).

- Nivel inferior (operativo): La información que se requiere en este nivel es la del estado de las actividades actuales dentro de las funciones básicas del negocio, gracias a esto se toman decisiones rutinarias, como por ejemplo una estimación de cantidad de materia prima requerida y abastecimientos.
- Nivel medio: Las decisiones tomadas en este nivel requieren de más tiempo y experiencia, esta implica revisión de resúmenes, análisis de datos históricos que permitan planificar, controlar operaciones e implementar políticas.
- Nivel superior: En el nivel superior se toman decisiones a largo plazo, generalmente impredecibles, relacionadas con actividades pasadas, presentes y futuras que requieren de una mayor experiencia. La información de este nivel encamina hacia la planificación estratégica, la asignación de recursos y la formulación de políticas.

2.9 Data Governance (DG)

Las empresas tienen continuamente que enfrentarse a la planificación de estrategias del negocio, el cumplimiento de los aspectos legales y modelos de negocio centrándose en el cliente, para apoyar el trabajo de estas existe DG que es utilizado para mejorar y mantener la calidad de los datos (Panian, 2010).

DG se preocupa de administrar las reglas de la gestión de los datos, en otras palabras, es la autoridad de los datos, lo que implica una cadena de mando y la delegación de funciones y responsabilidades. Más específicamente es un sistema de derechos de decisión y rendición de cuentas para los procesos relacionados con la información, ejecutados de acuerdo con lo acordado en modelos que describen las acciones a tomar con respecto a la información, y

cuándo, en qué circunstancias y mediante qué métodos para fomentar un comportamiento deseable en el uso de datos (Pennanen, 2014).

Los planes de DG fueron diseñados para ayudar a lograr claridad, asegurar el valor de los esfuerzos, crear una misión concisa, mantener enfoques claros y definir éxitos medibles en las organizaciones (Gwen, 2015). Los planes de DG se diferencian según el área de trabajo en la que se centraran, Thomas Gwen (Gwen, 2015) define 6 áreas de trabajo en un plan de DG estas pueden apreciarse en la Figura 2-5.

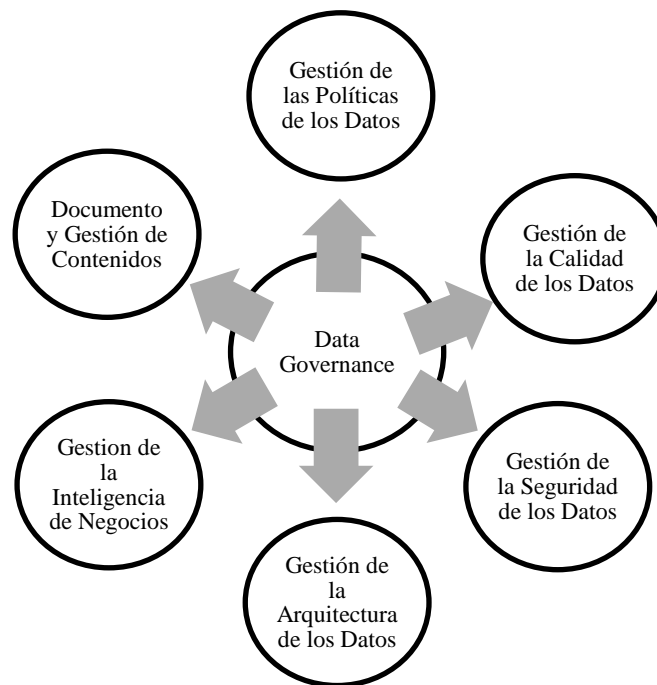


Figura 2-5: Data Governance (Fuente: Propia).

Las dos áreas de trabajo más importantes para este proyecto son la gestión de calidad de los datos y la gestión de seguridad de los datos, las cuales ya fueron definidas en las secciones 2.7 y 2.5 respectivamente.

2.10 Dirección de Servicios Informáticos

La Dirección de Servicios Informáticos (DSI) depende de la Vicerrectoría de Administración y Finanzas de la Universidad Católica de la Santísima Concepción y tiene como objetivo estratégico proveer y administrar servicios y tecnologías de información, adecuados a las necesidades de la comunidad universitaria, apoyando los procesos de enseñanza, investigación y extensión, así como apoyar la gestión académica administrativa, fomentando una cultura tecnológica en el uso eficiente y seguro de los recursos, en concordancia con los valores, principios y filosofía de nuestra Universidad (Dirección de Servicios Informáticos de la UCSC, 2016). Para ver catálogo de sistemas revisar Anexo A.

2.11 Sistema académico curricular (SIMBAD)

Sistema de la DSI encargado de registrar todos los procesos involucrados y la información relacionada con el ciclo completo del alumno o académico en su paso o permanencia por la Institución (Dirección de Servicios Informáticos de la UCSC, 2016).

Sus principales funcionalidades son:

- Admisión del Alumno.
- Formalización de Matrícula.
- Control de Pago y Aranceles.
- Control Registro de Programas de Estudio.
- Control y Administración de Cursos y Actividades Docentes.
- Registro y Procesamiento de Inscripción de Cursos.
- Control, Mantención y Seguimiento del estado curricular de los Alumnos.
- Registro Docente de Académicos.
- Control del flujo de los documentos y estados de los principales procesos o trámites.

En la Tabla 2-2 se pueden apreciar los módulos que están disponibles en SIMBAD

Tabla 2-2: Principales módulos de SIMBAD (Fuente: Propia).

MÓDULOS DE SIMBAD	
Módulo	Descripción
Admisión	Registro, control y procesos involucrados en la recepción y resolución de solicitudes de ingreso a la Universidad (autónoma y DEMRE).
Registro Académico	Permite el control de los procesos involucrados en la situación curricular de los alumnos; mantiene la información curricular de ellos (inscripción, solicitudes, notas, estado académico). Es responsable del registro y control de actividades docentes (cursos, profesores, salas, certificación, resoluciones, normas reglamentarias, indicadores, etc.), además de la mantención del registro de personal docente, control y mantención de planes de estudio, programación académica.
Título y Grado	Registro, control y seguimiento de actividad de titulación, control expediente de titulación y BD de egresado y titulados.
Matrícula	Registro y control de la realización de los trámites necesarios para la formalización de la matrícula de los estudiantes de 1er año y anteriores: compromiso económico, validación académica, documentación, aplicación de beneficios.
Bienestar	Registro y asignación de solicitudes de beneficio económico o social, antecedentes socioeconómicos, pase escolar, crédito fiscal y/o becas internas o externas.
Control Financiero	Permite el control y registro de los compromisos de pago y la recaudación asociada a los aranceles de matrícula: cuenta corriente, inventario de documentos, cobranza y caja.

Capítulo 3 . Estado del arte

Para contextualizar el presente proyecto, se han consultado diversas fuentes de información, dentro de las cuales se encontraron proyectos de similares características, además de documentos que manejan temas de interés acordes a los objetivos perseguidos en el desarrollo de este proyecto.

En el pasado las organizaciones no tenían mayor preocupación por los datos, ya que estos solo provenían de fuentes propias, en cambio en la actualidad los datos recogidos y analizados han superado estos alcances pues provienen de distintas fuentes, siendo la calidad de estas un problema. La calidad de los datos no depende sólo de sus propias características, sino también del entorno empresarial de los datos, incluyendo los procesos de negocio y los usuarios (Cai & Zhu, 2015).

3.1 Gestión de la seguridad de la información

En la actualidad gran parte de las organizaciones utilizan sistemas de información, ya que gracias a estos pueden procesar, almacenar y transmitir datos digitales; por lo que no tomar resguardos en los datos puede generar grandes problemas de seguridad. Es primordial que la organización entera se encuentre familiarizada y adopte enfoques comunes en lo que respecta con la seguridad de los sistemas de información (Dhillon & Backhouse, 2000).

La gestión de la seguridad de los datos es muy importante para las organizaciones, ya que con esta pueden proteger datos confidenciales, además los consumidores cada vez solicitan un mayor acceso a los datos y a las aplicaciones, y con esto las amenazas de seguridad crecen (Nazareth & Choi, 2015).

3.1.1 Principios para una correcta seguridad de la información

Según Dhillon y Backhouse (Dhillon & Backhouse, 2000) los principios básicos que debe seguir la seguridad de la información son: confidencialidad, divulgación no autorizada, integridad, modificación no autorizada, disponibilidad y retención no autorizada. Se

proponen una serie de principios para proveer una correcta seguridad de la información, siendo estos los primeros pasos que debería seguir una organización:

- **Responsabilidad:** Es muy importante que los miembros de una organización comprendan muy bien cuáles son sus funciones y responsabilidades, para que así puedan desarrollar su trabajo de manera adecuada.
- **Integridad:** La integridad para los miembros de una organización es muy importante, ya que, la empresa posee información valiosa. Es de gran importancia preocuparse por mantener y preservar la integridad en la empresa.
- **Confianza:** Debe existir confianza ente los miembros de una organización, creyendo que actuarán en conformidad a las normas propuestas por la compañía.
- **Ética:** Los miembros de una organización deben actuar con ética, en cuanto a las normas, en los comportamientos informales y donde no existan reglas.

A pesar de que estos son principios claves para una buena gestión de seguridad de información, se tiene que poner énfasis, principalmente, en la forma en que se aplican.

3.1.2 Un marco para la seguridad de la información

Una cultura de la seguridad de la información se define como la manera en la que se hacen las cosas en la organización para proteger los activos de la información, para este propósito es que se propone un marco que considera componentes técnicos, de procedimientos y del comportamiento humano (Da Veiga & Eloff, 2010).

Según Robbins (Da Veiga & Eloff, 2010), el comportamiento organizacional se refiere a lo que hace una organización, cómo se hace y cuál es su desempeño.

Por tanto, el marco de seguridad de la información debiese enfocarse en 2 aspectos principalmente: la política y el comportamiento de los empleados. A continuación, se definen estos aspectos:

- i. **Influenciar el comportamiento y el cultivo de una cultura de la seguridad de la información:** Tiene por objetivo inculcar el comportamiento de la seguridad de la información, que esta conduzca a la protección de los activos, basados en políticas

y códigos de ética en cuanto a la seguridad de la información en las organizaciones y por lo tanto establecer una cultura en cuanto a seguridad.

- ii. Crear un marco para la cultura de la seguridad de la información: Tiene como objetivo estudiar el comportamiento actual de la seguridad de la información y posteriormente crear una cultura para ésta.

3.1.3 Un modelo dinámico para la seguridad de la información

La seguridad de la información es un tema que preocupa a todas las organizaciones, ya que existe una gran variedad de amenazas, los administradores pueden elegir como contrarrestarlos con varias estrategias, tales como; la detección, la disuasión, la reducción de vulnerabilidad, la educación, entre otros. Es difícil de estimar el valor a invertir en seguridad de información, ya que las amenazas son impredecibles, estas pueden afectar a la organización en efecto dominó, afectando finalmente la reputación de la organización (Nazareth & Choi, 2015).

Nazareth y Choi (Nazareth & Choi, 2015), presentan un modelo de simulación para la gestión de la seguridad de la información, el cual integra varias facetas de la seguridad, como lo son la vulnerabilidad del software, evaluación de riesgos, posibles ataques, la detección de amenazas, la disuasión, la seguridad y los costos. El modelo hace hincapié en un modelo cuantificable y fácilmente verificable, proporcionando directrices claras sobre la inversión en seguridad y su impacto, mostrando que, si los ataques al sistema no se controlan, la vulnerabilidad del sistema aumenta. Además, el modelo se plantea a nivel organizacional y es ensamblado en base a la percepción de la misma organización. Estableciendo planes de inversión para las distintas áreas de la seguridad de la información, mostrando los costos asociados a la realización de este y a la pérdida que pudiese provocar el no realizarlo (Nazareth & Choi, 2015).

3.2 Gestión de Data Governance

Como se ha descrito anteriormente Data Governance (DG) se preocupa de la gestión de las reglas de la gestión de los datos, en otras palabras, de la autoridad de los datos, lo que

implica una cadena de mando y la delegación de funciones y responsabilidades (Pennanen, 2014).

Al implementar DG, independientemente del enfoque con el que se aplique se espera que se logren los siguientes objetivos (Gwen, 2015):

- Activar una mejor toma de decisiones.
- Reducir los roces operativos.
- Proteger las necesidades de los datos.
- Adoptar enfoques comunes a los problemas de datos.
- Construir estándares y procesos repetibles.
- Reducir los costos y aumentar la eficacia a través la coordinación de los esfuerzos.
- Garantizar la transparencia de los procesos.

A continuación, se explicarán algunos de los enfoques propuestos para poder abordar y llevar a cabo la implementación de DG.

3.2.1 Un modelo para Data Governance

DG, como parte de la gestión de calidad de los datos, especifica el marco de los derechos de decisión y rendición de cuentas para fomentar un comportamiento deseable en el uso de datos. El modelo de DG está compuesto por roles de la gestión de calidad de los datos, las áreas de decisión y las principales actividades y responsabilidades (Wende, A Model for Data Governance – Organising Accountabilities for Data, 2007).

DG se compone de las funciones de gestión de calidad de datos, áreas de decisión, principales actividades y responsabilidades; estos tres componentes están dispuestos en una matriz como se muestran en la Tabla 3-1, donde las columnas indican los cargos en la gestión de calidad de datos, las filas las áreas de decisión y las actividades principales; la matriz se llena con las responsabilidades, es decir, especifican grados de autoridad entre las funciones, las áreas de decisión y responsabilidades. Este modelo tiene como ventaja que es una

configuración única para cada empresa, ya que cada una define de forma distinta la gestión de calidad de datos.

Tabla 3-1: Modelo de Data Governance (Wende, A Model for Data Governance – Organising Accountabilities for data quality management, 2007).

ROLES	Patrocinador ejecutivo	Consejero de Data Governance	Administrador de los datos	...
AREA DE DECISIÓN				
Planificar iniciativas de calidad de datos	A	R	C	
Planificar iniciativas de calidad de datos	I	A	R	
Definir procesos de producción de datos		A	R	
Definir roles y responsabilidades	A	R	C	
...				

Como se puede apreciar en la Tabla 3-1 las filas detallan las áreas de decisión que contempla el plan de DG y las columnas representan los cargos presentes en la organización. La tabla es llenada con 3 posibles letras, indicando los grados de autoridad dependiendo del cargo. La R indica responsable, la A indica segundo responsable y la I indica informante, siendo una configuración específica para cada organización.

3.2.2 Data Governance con enfoque ágil

Marty Moseley (Moseley, 2008) identificó que una manera mucho mejor de abordar el tema DG usando una metodología ágil, ya que gracias a este enfoque se pueden ver resultados de manera mucho más rápida. Implementado procesos y políticas de gobierno en trozos más pequeños de la organización se busca obtener un proceso de adaptación y aprendizaje más rápido luego de cada iteración.

En la Figura 3-1 se puede apreciar cómo se obtendrá el proceso de adaptación luego de cada iteración, realizando en primera instancia una auditoria del estado actual de los datos, para continuar con la clarificación o creación de las políticas necesarias para el manejo de los datos, para así aplicar los procesos necesarios para solucionar los problemas encontrados

en la auditoría y finalmente se realizarán controles donde se compararán los resultados obtenidos (Moseley, 2008).

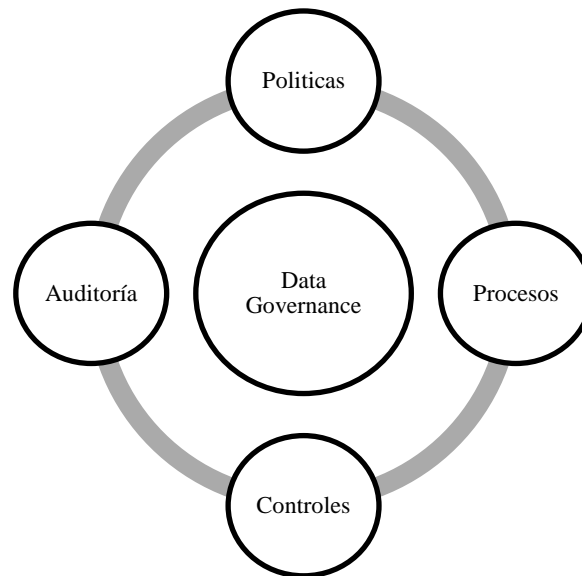


Figura 3-1: Data Governance ágil (Fuente: Propia).

El mayor beneficio que se obtiene al realizar un plan de DG con enfoque ágil, es que este ayuda a que los resultados obtenidos sean visibles en un periodo menor que si se aplicara de otra forma.

3.2.3 Una morfología para Data Governance

Distintos investigadores proponen variadas recomendaciones en cuanto a DG, como por ejemplo, diagramas funcionales que vinculan los derechos de decisión con roles (Dyché, 2007), la especificación de las fuentes de los administradores de datos (Economist Intelligence Unit, 2008), establecer derechos de decisión o recomendaciones para establecer DG en las estructuras organizativas de la empresa (Khatri & Brown, 2010). Todos estos enfoques se caracterizan por centrarse en un solo aspecto de DG, conduciendo a soluciones aisladas.

Ritchey (Ritchey, 2006) aborda el objetivo de capturar la totalidad de la organización con DG, proponiendo una morfología, que define la estructura y la disposición de las partes de un objeto, y cómo estos se ajustan a crear un todo. La morfología constituye una teoría

analítica, es decir, se identifican y estructuran de los conceptos básicos de la organización de gobierno de datos, dividiéndolos en 2 dimensiones, como muestra la Figura 3-2, la primera en los objetivos de la organización y la segunda en la estructura organizativa (Otto, 2011).

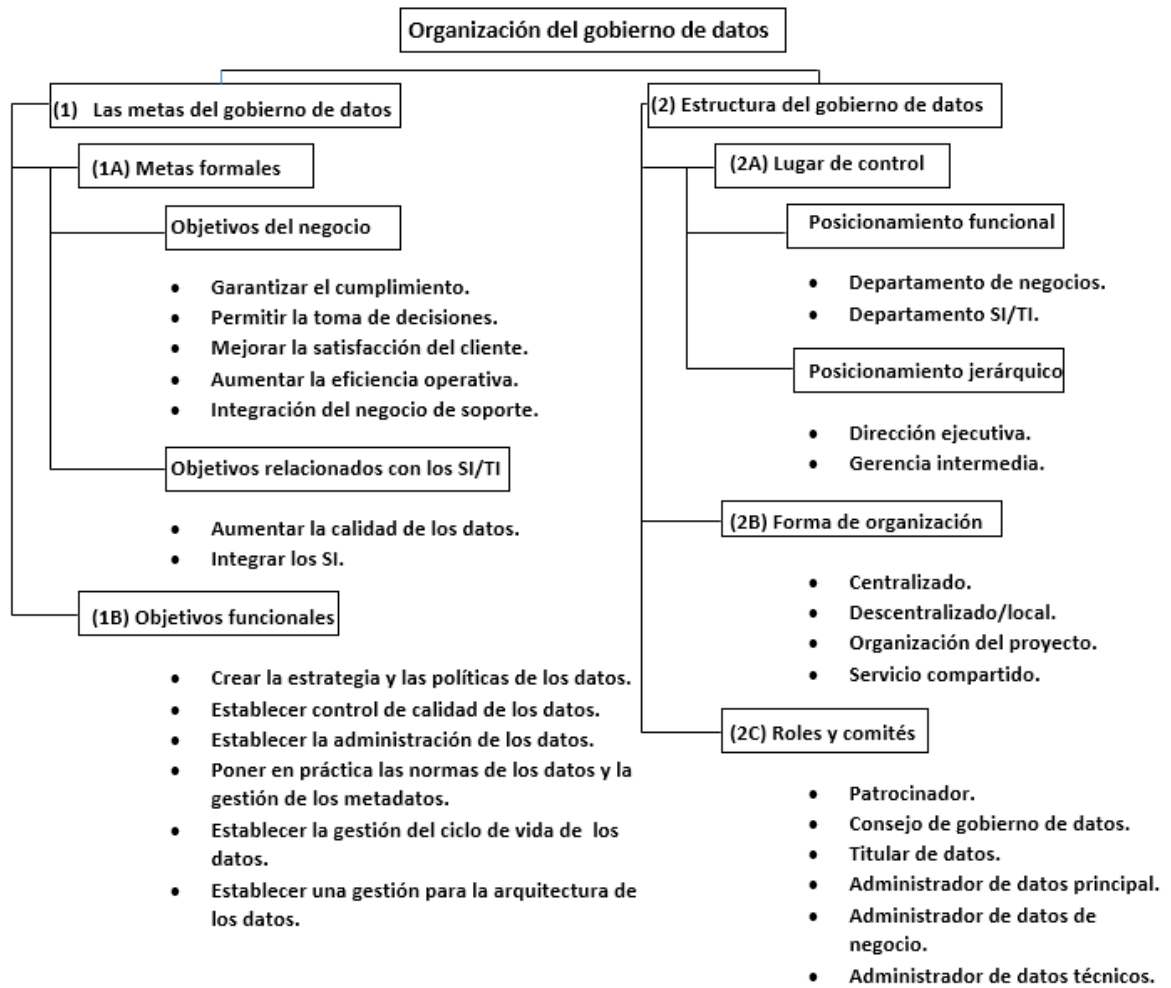


Figura 3-2: Una morfología para la organización de data governance (Otto, 2011).

Esta morfología sirve para estructurar el tema de la investigación de DG, permite identificar y evaluar los patrones de DG y el análisis detallado de los factores de contingencia (Otto, 2011).

3.2.4 Áreas de aplicación de Data Governance

A continuación, se presenta una serie de áreas donde ha sido aplicado de manera exitosa DG.

a. Industria de servicios financieros

Patrick Egan, de la Universidad de Regis (Patrick, 2011), realiza un caso de estudio para los servicios financieros, donde demuestra la importancia que tiene el aplicar DG en relación a la calidad de los datos para estas industrias.

Para su realización el investigador optó por una metodología de 3 fases:

- Fase I: Consiste en la revisión de la literatura existente, donde se investigó todo lo relacionado con los datos, trabajos anteriores en lo que es calidad de datos, gestión de datos maestros y gestión de la información.
- Fase II: Consiste en revisar la implementación, si esta cumplía con lo anteriormente investigado, donde se realizaron entrevistas a los encargados.
- Fase III: Consiste en enviar una encuesta a varios ejecutivos de la empresa donde se podía apreciar la importancia que ellos encontraban en los datos para la organización.

Uno de los primeros problemas que surgen en este estudio, es que las empresas no se dan cuenta del gran costo que implica el tener datos incorrectos, por lo que para ellos el tema de calidad de datos no es muy atractivo. Otro de los problemas es el crecimiento exponencial de los datos, ya que en muchos casos faltan políticas de retención, no existe un conocimiento acerca del dominio y un miedo a la destrucción de estos, incluso al haber cumplido ya la fecha de caducidad (Patrick, 2011).

El estudio logró identificar que entre un 40% y 50% de las empresas, existe una conciencia en lo que respecta calidad de los datos, ellas poseen tecnologías para esto, sin embargo, no tienen una estrategia empresarial coherente para su uso.

Las empresas que poseen un grado de madurez, han implementado DG, lo han hecho de manera adecuada con la administración, con el presupuesto y el patrocinio. En las decisiones operativas y estratégicas los datos con lo que se trabaja son de confianza.

b. Data Governance en la medicina

Gil y Colomer (Gil & Colomer, 2014) se dieron cuenta que en el sector general de salud los sistemas de información están teniendo un gran impacto en la actualidad, tanto para el área médica, como para las de análisis de datos, historial médico, análisis clínicos, gestión en centros de salud, la administración hospitalaria entre otros. Identificaron la importancia de mantener una buena calidad de los datos y los riesgos asociados al mal uso de estos.

Los investigadores decidieron realizar un plan de DG estableciendo un marco en el cual se incluyeran todas las funciones, donde el mayor objetivo fue mejorar la accesibilidad a los datos. Para poner en marcha DG fue necesario que la organización definiera su estrategia en cuanto a la planificación y al control.

Actualmente DG es considerado como esencial para el centro sanitario para la toma de decisiones relacionadas con el cuidado de la salud de las personas. Además de proporcionar mayor control sobre la creación, manejo, mantenimiento, almacenamiento, uso e intercambio de información vital para el negocio (Gil & Colomer, 2014).

c. Data Governance en el control de riesgos

Mosley (Moseley, 2008) se percató que gracias a los avances y la tecnología se han podido reunir y almacenar gran cantidad de datos, pero para marcar la diferencia, tienen que ser accesados por la persona correcta de manera inmediata generando una verdadera ventaja competitiva, lo que puede ocasionar problemas de seguridad. Por lo que surgió la necesidad de desarrollar estrategias y programas de DG comprendiendo personas, procesos y tecnologías necesarias para una buena gestión de los datos, para así poder manejar correctamente los riesgos que estos podrían ocasionar. Identificó que los puntos más importantes a medir eran la confidencialidad, integridad y disponibilidad de los datos.

El investigador propuso que la mejor manera de enfrentar un plan de DG era dividirlo en 2 fases, la primera para las capas de herramientas y soluciones con el fin de que protejan la integridad y seguridad de los datos, y la segunda para las soluciones de vigilancia e

información con el fin de ayudar a las empresas a garantizar que todo el mundo cumpla con las políticas de control de datos de la organización (Moseley, 2008).

DG es un proceso continuo circular que implica a todos, a un proceso de control, auditoría y evaluación de los datos, el cual busca que los usuarios puedan comprender mejor el uso y el riesgo de los datos, para garantizar que se le está sacando el mayor provecho y uso a los datos (Moghe, 2009).

3.2.5 Dificultades al aplicar Data Governance

Varios problemas surgen en la actualidad, algunos de ellos son:

1. Mala calidad de los datos internos provocando desastrosas tomas de decisiones y por ende malas estrategias de negocios (Moseley, 2008).
2. No contar con disponibilidad de recursos (Power Data. Especialista en gestión de datos, 2016).
3. No cumplir con las reglas establecidas en los procesos (Power Data. Especialista en gestión de datos, 2016).
4. Establecer objetivos irrealistas (Power Data. Especialista en gestión de datos, 2016).
5. No tener claro el concepto de DG, creyendo que no solo es la elaboración de políticas y derechos de decisión sino que también la ejecución de las estas (SAS Latín América, 2016).

3.2.6 ¿Por qué usar Data Governance?, principales resultados

DG es de gran ayuda a las organizaciones ya que esta busca lograr (Gwen, 2015):

- Claridad
- Asegurar valor en los esfuerzos
- Crear una misión clara
- Mantener alcance y enfoque
- Establecer responsabilidades
- Definir éxitos medibles

Y gracias a la implementación de esto, la organización puede (Gwen, 2015):

- Activar una mejor toma de decisiones
- Proteger las necesidades de los datos
- Que los involucrados adopten enfoques comunes a los problemas de datos
- Construir estándar
- Reducir los costes y aumentar la eficacia a través la coordinación de los esfuerzos
- Garantizar la transparencia de los procesos

Un modelo de DG ayuda a las empresas en la estructuración de sus responsabilidades de calidad de datos, en relación a las funciones propuestas, áreas de decisión y responsabilidades. Pueden utilizar el modelo de DG como un dispositivo de comunicación de toda la compañía, para los roles de la gestión de calidad de datos y su tipo de interacción con las actividades y decisiones específicas. La implementación de DG ayuda a la empresa a encontrar una configuración que mejor se adapte a esta, y, por lo tanto, maximizar la contribución positiva de la gestión de calidad de datos a los objetivos generales de la empresa de una compañía. Dependiendo del nivel de granularidad, una empresa podría incluso definir más de un modelo de gobierno de datos. Por ejemplo, se puede definir un modelo por área de decisión o un modelo para el nivel corporativo y un modelo adicional por unidad de negocio (Wende, A Model for Data Governance – Organising Accountabilities for data quality management, 2007).

3.3 Conclusión y discusión del capítulo

Gracias a lo investigado se pudo conocer, en cuanto a seguridad de los datos, que la mayoría de los autores se preocupan de evaluar la seguridad de la información, sin embargo, los datos son la materia prima de la información por lo que los principios utilizados son los mismos, en este proyecto se buscó detectar problemas de seguridad en los datos, crear procesos y políticas en base a las normas descritas por la ISO27001, ya que se adaptan en su mayoría a lo propuesto por los distintos autores estudiados, con el fin de controlar el acceso a los datos, educar y clarificar las normativas para el uso de los datos a los usuarios del sistema.

En cuanto al plan de DG enfocado en la calidad de los datos luego de lo estudiado, se pudo concluir que a pesar de que es aplicable a todo tipo de organizaciones, no existe un enfoque común para estos, ya que cada organización tiene que establecer sus objetivos, sin embargo, lo más aconsejable es realizar el plan de DG en una parte de la organización en primera instancia, para así poder lograr ver resultados de manera más rápida.

En este proyecto se realizó un plan de DG enfocado en la calidad y seguridad de los datos donde se buscó proteger los datos, clarificar las responsabilidades y políticas; y asegurar una buena calidad de los datos, para esto en primera instancia se procedió a diagnosticar el estado actual en cuanto a calidad y seguridad de los datos del sistema SIMBAD y para eso se utilizó como ejemplo la metodología utilizada por Patrick, descrita en la sección 3.2.3 punto a, evaluando las dimensiones descritas en el capítulo 2 sección 2.4.

A modo de resumen y comparación de este capítulo se presenta la Tabla 3-2 y la Tabla 3-3.

Tabla 3-2: Comparación de los enfoques de la gestión de seguridad de la información (Fuente: Propia).

COMPARACIÓN DE LOS ENFOQUES DE LA GESTION DE SEGURIDAD DE LA INFORMACIÓN		
Enfoques	Atributos a medir	Beneficios
Principios fundamentales para la seguridad de la información	<ul style="list-style-type: none"> -Disponibilidad. -Confidencialidad. -Integridad. -Responsabilidad. -Confianza. -Ética. 	<ul style="list-style-type: none"> -Gran confianza en los participantes de la organización. -Genera lazos con la organización.
Marco para la seguridad de la información	<ul style="list-style-type: none"> -Se enfoca en la política y en el comportamiento de los empleados. 	<ul style="list-style-type: none"> -Cultiva una cultura de seguridad en la organización. -Se trabaja en las individualidades de las personas.
Un modelo dinámico	<ul style="list-style-type: none"> Integra facetas de la seguridad: - vulnerabilidad del software. -Evaluación de riesgos. -Posibles ataques. -La detección de amenazas. -La disuasión, la seguridad. -Los costos. 	<ul style="list-style-type: none"> -Muestra directrices claras de la inversión a realizar en seguridad. -Muestra el impacto que puede provocar amenazas al sistema.

Tabla 3-3: Comparación de los enfoques Data Governance (Fuente: Propia).

COMPARACIÓN DE LOS ENFOQUES DATA GOVERNANCE		
Enfoques	Características	Beneficios
Data Governance Ágil	-Se implementan procesos y políticas de gobierno en trozos pequeños de la organización, generando diversas iteraciones.	- Se aprecian resultados a un corto plazo. - se genera un proceso de adaptación y aprendizaje luego de cada iteración.
Modelo de Matriz	-Se crea una matriz donde las columnas indican los papeles en la gestión de calidad de datos, las filas, áreas de decisión y las actividades principales. -La matriz se llena con las responsabilidades especificando grados de autoridad.	-Se estructuran responsabilidades de calidad de datos. - Configuración única para cada empresa. -Define claramente los roles y la decisiones específicas, mejorando la comunicación en la compañía.
Morfología Analítica	-La morfología constituye una teoría analítica que identifica y estructura los conceptos básicos de la organización de DG. --Se centra en 2 grandes áreas: -> Objetivos de la organización ->Estructura organizativa	-DG captura la totalidad de la organización. - Estructurar el tema de la investigación de DG -permite identificar y evaluar los patrones de DG -Entrega un análisis detallado de los factores de contingencia

Por otra parte, se presentan las principales consecuencias de implementar DG en distintos sectores, a modo de comparación se presenta la Tabla 3-4.

Tabla 3-4: Resumen de implementaciones (Fuente: Propia).

RESUMEN DE IMPLEMENTACIONES		
Se ha aplicado	Problemas	Mejoras
Industria de Servicios Financieros	-Falta de políticas en relación a los datos -No existe aprecio por la calidad de los datos. -Miedo a la destrucción.	-Se cuantifico el costo de la mala calidad de los datos. - Se crearon políticas en relación a los datos. -Se establecieron parámetros de caducidad de los datos.
Control de riesgos para las empresas	-Presencia de riegos en el acceso a los datos -Tiempo de acceso a los datos -Problemas de seguridad	-Se mejoró integridad y seguridad de los datos e información. - Se crearon políticas de control de datos que ordenaban la organización.
En la medicina	-Inconsistencia en los datos. -Mal acceso a los datos	-Creación de estándares y políticas solucionando problemas de inconsistencias. -Mayor control sobre la creación, manejo, mantenimiento, almacenamiento, uso e intercambio de información

Capítulo 4 . Diagnóstico de calidad y seguridad de los datos

En el presente capítulo se presenta el diagnóstico en relación a la calidad y seguridad de los datos del sistema SIMBAD, para la realización de este fue necesario la realización de entrevistas a los encargados del funcionamiento del sistema, la formulación de encuestas a distintos usuarios del sistema; y además una revisión al sistema (documentación diseño e interfaz).

4.1 Entrevistas

Para determinar en qué sistema se trabajó, que ámbitos se evaluaron y para diagnosticar, fue necesario la realización de una serie de entrevistas a lo largo de todo el proceso.

Primera entrevista

- Objetivo: Conocer mayormente el funcionamiento del sistema SIMBAD y conseguir permisos para desarrollar el proyecto en dicho sistema.
- Entrevistado: Cristian Cárdenas.
- Conclusión entrevista: Se pudo concluir que los datos almacenados en el sistema SIMBAD son de gran importancia para la universidad, ya que en su mayoría son datos confidenciales de los alumnos y de los académicos, por lo que asegurar la seguridad de los datos es de vital importancia, además este alimenta a la plataforma de inteligencia de negocios CUBIX la cual entrega gran información a la universidad, por lo que asegurar una buena calidad de los datos en el sistema SIMBAD es muy importante. Se solicitaron los permisos pertinentes y se quedó en espera de una pronta respuesta.

Segunda entrevista

- **Objetivo:** Conocer el funcionamiento del sistema en cuanto a programación, recaudar información relevante para el proyecto.
- **Entrevistado:** José Comas.
- **Conclusión:** Se concluyó que el sistema en su mayoría es manejado por sesiones de la base de datos de Oracle, además se consiguió la especificación de casos de uso por módulos y el modelo de la base de datos del sistema SIMBAD.

Cabe destacar que durante todo el proceso se mantuvo contacto con el señor José Comas, quien fue de gran ayuda para la realización de este trabajo, solicitando accesos para la revisión al sistema y ayudando con información relevante para el proyecto, gracias a esto se pudo realizar un diagnóstico actual para los datos del sistema SIMBAD.

4.2 Encuesta de calidad de los datos

Como se pudo concluir de las entrevistas realizadas en el punto 4.1, asegurar una buena calidad de datos en el sistema SIMBAD es de suma importancia, ya que alberga una gran cantidad de datos que son de vital importancia para el correcto funcionamiento de la universidad.

Para la realización del diagnóstico de calidad de los datos fue necesario realizar una evaluación previa, para esto se realizaron encuestas, las cuales fueron aplicadas a distintos usuarios de SIMBAD, en las que se evaluaron las 6 dimensiones definidas en el capítulo anterior (completitud, consistencia, exactitud, unicidad, oportunidad y validez), realizando entre 2 y 3 preguntas por cada atributo.

El principal objetivo de la realización de encuestas fue conocer la opinión directa de los distintos usuarios y su apreciación y evaluación en cada dimensión. La encuesta fue aplicada a 6 usuarios, entre ellos jefes de carrera, secretarías y administrativos, donde además de las preguntas genéricas de la encuesta, se les consultó las distintas funciones que realizaban con el sistema, las respuestas obtenidas se encuentran detalladas en el Anexo C.

Ya que el formato de respuesta de la encuesta era “SI” ”NO” ”NO APLICA”, se calculó el porcentaje de respuestas correctas por dimensión de la siguiente forma (Covella, 2005):

$$R = \frac{NRE}{NEN} * 100$$

Donde:

NRE: Número de respuestas esperadas.

NEN: Número de encuestados.

Para evaluar los resultados obtenidos de las encuestas se utilizó una rúbrica de 3 posibles criterios:

Excelente \longrightarrow Respuestas no esperadas < 3%

Aceptable \longrightarrow 2 % < Respuestas no esperadas < 29%

Insuficiente \longrightarrow Respuestas no esperadas > 28%

A continuación, se presentan las preguntas realizadas por dimensión y la respuesta esperada por pregunta en negrita:

- Completitud

1. ¿Al consultar los datos de un alumno o a un docente en el sistema, se presentan campos vacíos? **No**
2. ¿Al realizar ingresos en el sistema, ya sean de solicitudes, datos de académicos o alumnos, existen campos de llenado no obligatorios? **No**
3. ¿Se presentan registros con nombre o direcciones cortadas en el sistema? (Nombre: María Ignacia Jav) **No**

- Consistencia

1. ¿Al consultar por los datos de un docente, estos se encuentran siempre actualizados? **Si**

2. ¿Los datos de los alumnos son consistentes? Ejemplo (País de residencia: China) **Si**
 3. ¿Los datos registrados de los docentes en el sistema, son concordantes con los datos que posee la Dirección de Gestión de Personal? **Si**
- Exactitud
 1. ¿Al consultar distintos registros en el sistema, estos son desplegados en el sistema con el mismo formato? **Si**
 2. ¿Es posible registrar en el sistema a un alumno con año de nacimiento 1900? **No**
 3. ¿Al ingresar datos al sistema, independientemente del formato en que los ingrese, el sistema los corrige estandarizándolos todos? **Si**
 - Unicidad
 1. ¿Al consultar por un alumno en el sistema, es desplegado más de un registro? **No**
 2. ¿Es posible generar una reserva de sala en el sistema en un aula ya reservada? **No**
 3. ¿Al consular en el sistema existen el mismo número de matrículas que de registros de alumnos? **Si**
 - Oportunidad
 1. ¿Los alumnos al realizar actualización de los datos en DARA, estos son actualizados en un periodo máximo de 2 días hábiles en el sistema? **Si**
 2. ¿Las notas de examen de título son registradas en el sistema en no más de 2 días hábiles? **Si**
 - Validez
 1. ¿Puede ingresar datos en el sistema como la fecha o el Rut en distintos formatos (dd/mm/aaaa o dd-mm-aaaa)? **Si**
 2. ¿Se le permite ingresar al sistema Rut con más dígitos de los existentes (111.111.111-1)? **No**

3. ¿En el registro de personas en el sistema, campos como la dirección y el nombre son validados, evitando que se cometan errores de tipeo? **Si**

A continuación, se presentan los resultados obtenidos por dimensión, para revisar los resultados más detalladamente revisar el Anexo C.

- **Completitud:** En cuanto a la completitud, los resultados obtenidos como se pueden apreciar en el Gráfico 4-1 indican que un 39% de las respuestas de los encuestados no eran las esperadas, *siendo la principal causa que al realizar ingresos al sistema existen campos de llenados no obligatorios por ende al consultar por algún registro en el sistema se presentan campos vacíos.*

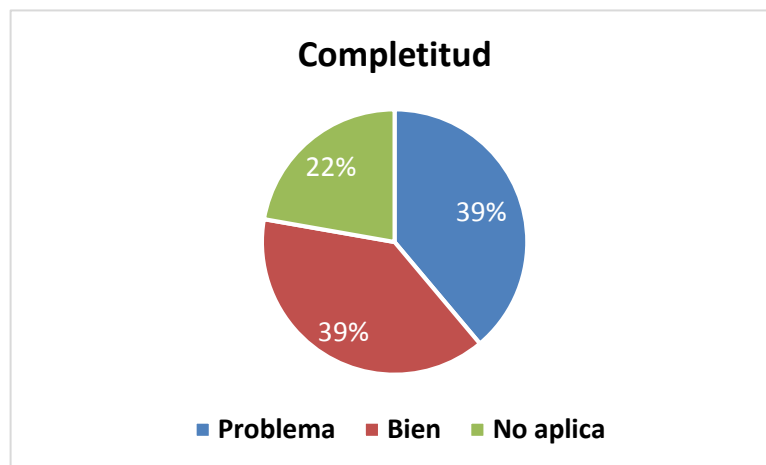


Gráfico 4-1: Resultado de la encuesta en completitud de los datos (Fuente: Propia).

- **Consistencia:** En cuanto a la consistencia, los resultados obtenidos como se pueden apreciar en el Gráfico 4-2 indican que en un 39% de las respuestas de los encuestados no eran las esperadas, *siendo la principal causa que no siempre se encuentran actualizados los datos, esto puede deberse a la fuente de donde provienen los datos, o que tan seguido el sistema realiza las actualizaciones de estos.*

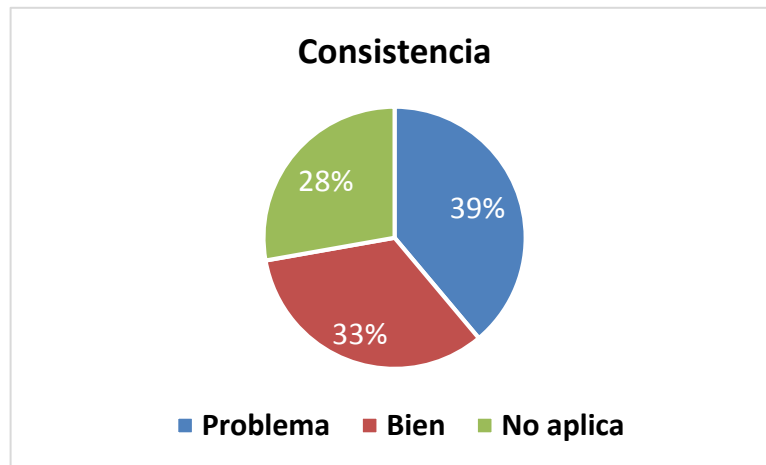


Gráfico 4-2: Resultado de la encuesta en consistencia de los datos (Fuente: Propia).

- Exactitud: En cuanto a la exactitud, los resultados obtenidos como se pueden apreciar en el Gráfico 4-3 indican que en un 22% de las respuestas de los encuestados no eran las esperadas, *siendo la principal causa la estandarización de los datos, lo que quiere decir que el sistema independiente del formato y fuente en el que se ingresen registros al sistema, este no es capaz de estandarizar todo lo ingresado a un solo formato.*

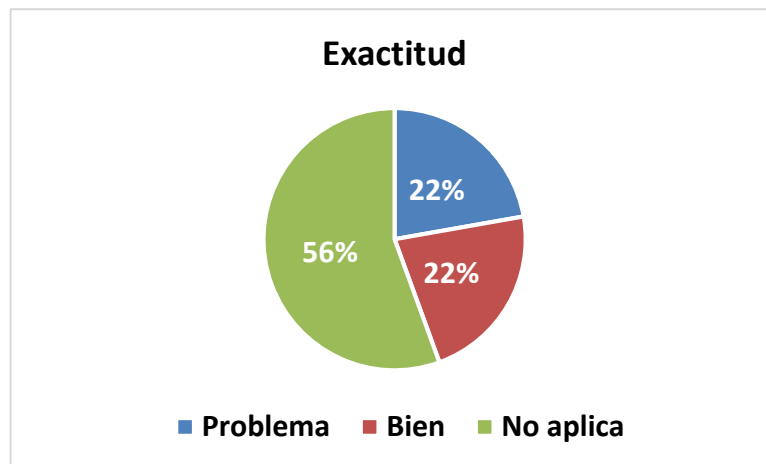


Gráfico 4-3: Resultado de la encuesta en exactitud de los datos (Fuente: Propia).

- Unicidad: En cuanto a la unicidad, los resultados obtenidos como se pueden apreciar en el Gráfico 4-4 indican que en un 11% de las respuestas de los encuestados no eran

las esperadas, *algunos usuarios consideraron que en algunos casos se ha presentado más de un registro o registros contradictorios al consultar datos en el sistema SIMBAD.*

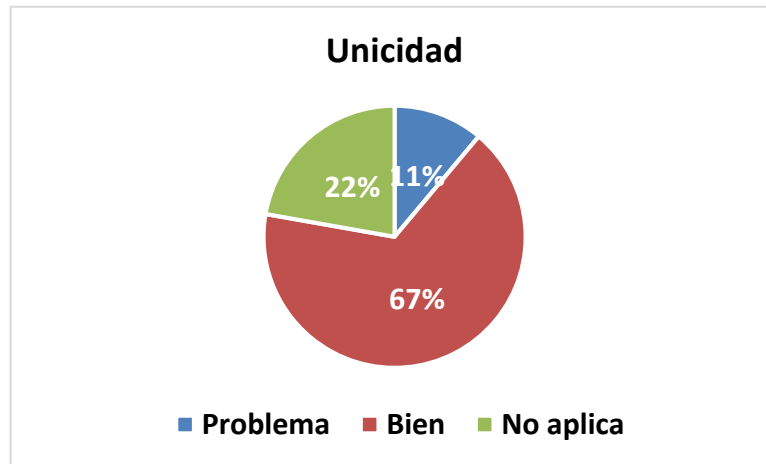


Gráfico 4-4: Resultado de la encuesta en unicidad de los datos (Fuente: Propia).

- Oportunidad: En cuanto a la oportunidad, los resultados obtenidos como se pueden apreciar en el Gráfico 4-5 indican que el 67% de los encuestados no se sienten capacitados para dar una opinión de las preguntas propuestas. Sin embargo, el 33% de los usuarios restantes opinaron que no existe problema alguno, *asegurando que las solicitudes en el sistema o nuevos registros son registradas en el sistema en un periodo máximo de 2 días hábiles, el cual es considerado como un tiempo aceptado en los parámetros de calidad.*

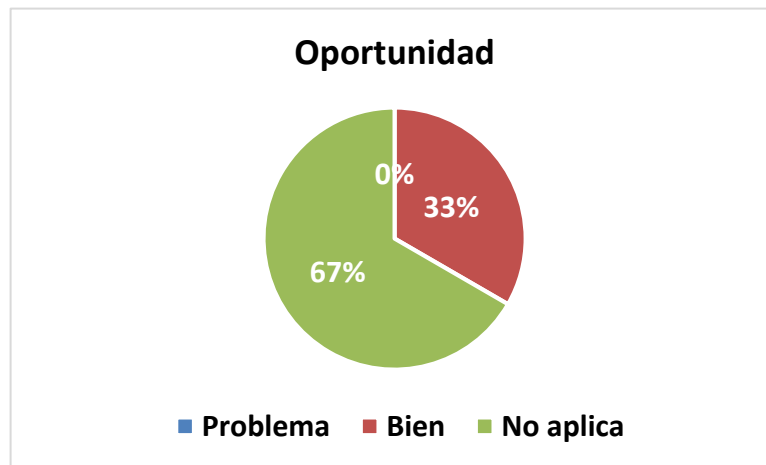


Gráfico 4-5: Resultado de la encuesta en oportunidad de los datos (Fuente: Propia).

- Validez: En cuanto a la validez, los resultados obtenidos como se pueden apreciar en el Gráfico 4-6 indican que en un 17% de las respuestas de los encuestados no eran las esperadas, *siendo la principal causa la no validación de los datos del sistema en lo que respecta nombres y direcciones.*

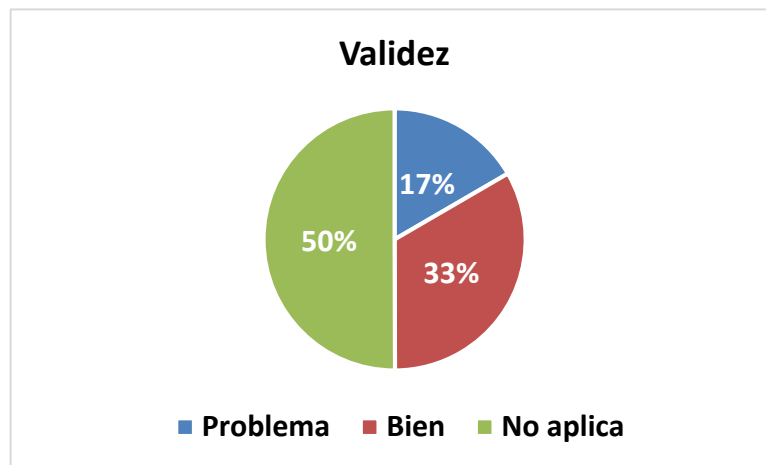


Gráfico 4-6: Resultado de la encuesta en validez de los datos (Fuente: propia).

Los resultados obtenidos, como se pueden apreciar en el Gráfico 4-7, presentan que las dimensiones de completitud y consistencia se encuentran en niveles de insuficiencia, esto puede deberse a que, en completitud, al momento de realizar registros en el sistema se presentan gran cantidad de campos de llenado no obligatorios, siendo la principal causa de que al consultar por algún registro se presenten gran cantidad de campos vacíos. Por otra parte, en la dimensión de consistencia uno de los principales problemas encontrados por los usuarios es que los datos no se encuentran siempre actualizados, esto puede deberse a la fuente de donde provienen los datos o que tan seguido el sistema es actualizado.

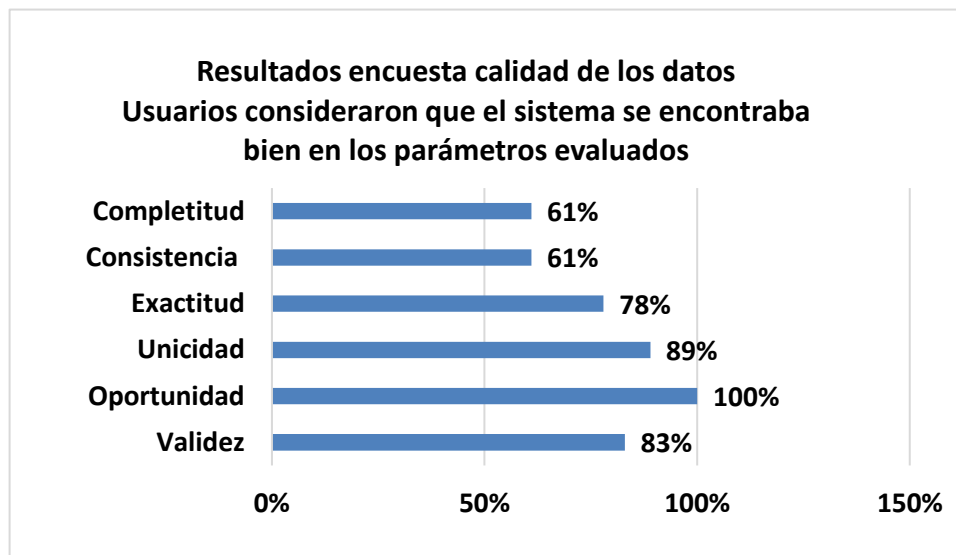


Gráfico 4-7: Resultados encuesta calidad de los datos (Fuente: propia).

4.3 Revisión del sistema en cuanto a calidad de los datos

Se realizó una revisión a la base de datos del sistema SIMBAD, con el fin de inspeccionar las tablas que juegan roles fundamentales en el área que administra la DARA. Para esto, como se muestra en la Tabla 4-1, se realizó un seguimiento de los posibles datos que son ingresados, consultados, eliminados o modificados en distintas tablas de la base de datos del sistema. Con la revisión se obtuvo un diagnóstico más claro del estado actual del sistema y se evaluaron aspectos que no pudieron ser revisados por medio de encuestas a los usuarios del sistema.

Con el fin de medir la consistencia de la base de datos del sistema es que se procedió a inspeccionar 60 tablas de la base de datos que se considera que juegan roles fundamentales en el funcionamiento del sistema SIMBAD, se realizaron distintas pruebas con el fin de medir las dimensiones propuestas en el capítulo anterior (completitud, consistencia exactitud, unicidad, oportunidad y validez). Cada tabla fue analizada con 3 roles distintos: el de secretarías, jefes de carrera y el de administrativos.

Se analizaron distintos aspectos, de los cuales se calculó un porcentaje de error, el cual fue calculado de la siguiente forma (Covella, 2005):

$$R = \frac{TPE}{TP} * 100\%$$

Donde:

TPE: Total pruebas con error.

TP: Total de pruebas.

Para evaluar los resultados obtenidos de la revisión se utilizó una rúbrica de 3 posibles criterios:

Excelente \longrightarrow Errores < 3%

Aceptable \longrightarrow 2 % < Errores < 29%

Insuficiente \longrightarrow Errores >28%

Tabla 4-1: Revisión a la base de datos de SIMBAD (Fuente: Propia).

CALIDAD DE LOS DATOS					
Dimensión	Tipos de errores	Número de tablas analizadas	Número de columnas analizadas	Número de columnas con errores	Porcentaje de error
Compleitud	Campos vacíos	60	178	25	14%
	Datos incompletos	60	178	4	2.2%
Consistencia	Datos irreales	50	136	6	4.4%
	Actualización tardía	60	178	0	0%
Exactitud	Distintos formatos	60	178	9	5.1%
	Límite de parámetros irreales	50	136	10	7.4%
	No existe estandarización	60	178	14	7.9%
Unicidad	Duplicidad	60	178	1	0.6%
Validez	Campos no se encuentran restringidos apropiadamente	50	136	10	7.4%
	Datos no son validados	50	136	14	10.3%

Los resultados obtenidos, como se pueden apreciar en el Grafico 4-8, muestran que, en aspectos de calidad de datos, el sistema SIMBAD se encuentra con buenos estándares, ya que, en las dimensiones analizadas, unicidad es la mejor evaluada y esta se encuentra casi en un 100% de perfección con lo que cae en el intervalo de excelente, a pesar de que los otros atributos no alcanzan a pertenecer a aquel intervalo, se encuentra en aceptable, pero con índices muy cercanos a la excelencia.

La dimensión más baja fue la exactitud, la cual obtuvo un 20,4% de error, ubicándose en el intervalo de aceptable, en esta los mayores problemas encontrados fueron los límites de los parámetros, como ejemplo se puede apreciar en la Tabla 4-2, donde la variable PERS_F_NACIMIENTO que a pesar de ser de tipo DATE, no se encuentra validado que el año de nacimiento ingresado sea correcto. Este tipo de problemas se registró en todas las variables de tipo DATE. Por otro lado, no existe estandarización en cuanto a mayúsculas y minúsculas, encontrándose en distintos formatos los datos.

Existe un gran número de tablas que poseen columnas vacías en el sistema, las cuales afectan directamente el porcentaje de completitud obtenido de los datos, por otra parte, los registros vacíos solo ocupan memoria, claramente estos no son utilizados recurrentemente por lo que se puede deducir que no son realmente necesarios para el funcionamiento del sistema.

En el siguiente capítulo se listan una serie de procedimientos que van en pro de mejorar la calidad de los datos del sistema SIMBAD.

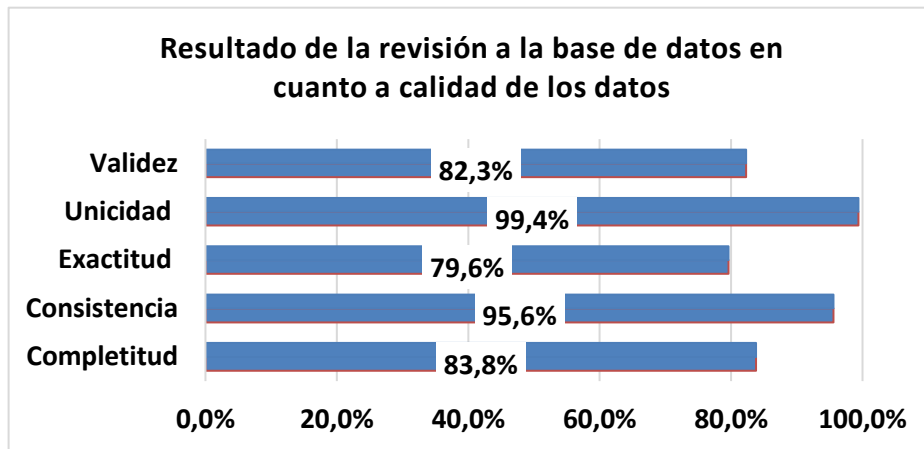


Gráfico 4-8: Resultados revisión base de datos de SIMBAD (Fuente: Propia).

Tabla 4-2: Tabla PERSONAS (Dirección de Servicios Informaticos, 2007).

TABLA PERSONAS	
PERS_ALIAS	VARCHAR2(50)
PERS_AP_MATERNO	VARCHAR2(20)
PERS_AP_PATERNO	VARCHAR2(30)
PERS_COMENTARIO	VARCHAR2(100)
PERS_C_EST_C	NUMBER <fk1>
PERS_DV_RUT	VARCHAR2(1)
PERS_E_MAIL	VARCHAR2(50)
PERS_F_NACIMIENTO	DATE
PERS_NACIONLD	NUMBER <fk2>
PERS_NOM1	VARCHAR2(20)
PERS_NOM2	VARCHAR2(40)
PERS_PASSWORD	VARCHAR2(15)
PERS_RUT	NUMBER <pk>
PERS_SEXO	VARCHAR2(1)
PERS_SUB_TIPO	VARCHAR2(4)
USERNAME	VARCHAR2(15)

4.4 Encuesta de seguridad de los datos

Como se pudo concluir en las entrevistas realizadas en el punto 4.1, es de suma importancia asegurar una correcta seguridad de los datos para el sistema SIMBAD, ya que en este se encuentran todos los registros académicos y datos personales de cada alumno o académico de la universidad.

Para la realización del diagnóstico de seguridad de los datos fue necesario realizar una evaluación previa, a través de una encuesta, la cual fue aplicada a distintos usuarios de SIMBAD, en las que se evaluaron las 3 dimensiones mencionadas en el capítulo anterior (confidencialidad, integridad y disponibilidad), realizando entre 2 y 3 preguntas por cada atributo.

El principal objetivo de la realización de la encuesta fue conocer la opinión directa de los distintos usuarios acerca de cómo evaluaban ellos cada dimensión propuesta. La encuesta fue aplicada a 6 usuarios, entre jefes de carrera, secretarías y administrativos, donde además de las preguntas genéricas de la encuesta, se les consultó las distintas funciones que realizaban con el sistema, las respuestas obtenidas se encuentran detalladas en el Anexo C.

Ya que el formato de respuesta de la encuesta era “SI” ”NO” ”NO APLICA”, se calculó el porcentaje de respuestas correctas por dimensión de la siguiente forma (Covella, 2005):

$$R = \frac{NRE}{NEN} * 100\%$$

Donde:

NRE: Número de respuestas esperadas.

NEN: Número de encuestados.

Para evaluar los resultados obtenidos de las encuestas se utilizó una rúbrica de 3 posibles criterios:

Excelente \longrightarrow Respuestas no esperadas < 3%

Aceptable \longrightarrow 2 % < Respuestas no esperadas < 29%

Insuficiente \longrightarrow Respuestas no esperadas > 28%

Las preguntas realizadas en la encuesta fueron escogidas de manera que fueran capaces de medir de manera clara las dimensiones y que a su vez no fueran engorrosas ni tediosas para los usuarios encuestados. Los resultados obtenidos en esta encuesta dieron a conocer una visión clara a los problemas presente de seguridad de los datos que los usuarios consideraban presentes en el sistema, para lo cual en el siguiente capítulo se definieron y esclarecieron procesos que ayudan a mejorar los problemas encontrados. A continuación, se presentan las preguntas realizadas por dimensión y la respuesta esperada por pregunta en negrita:

- Confidencialidad

1. ¿Firmó un acuerdo de confidencialidad, para el acceso en el sistema, de datos personales y confidenciales? **Si**
2. ¿Puede acceder a datos en el sistema que le son innecesarios para su trabajo? **No**
3. ¿Tiene conocimiento sobre políticas o normativas para la seguridad de los datos en el sistema impuestas por la universidad? **Si**

- Integridad

1. ¿Puede modificar libremente los datos personales de un estudiante o académico sin una solicitud previa del individuo en cuestión? **No**
2. ¿Puede eliminar registros del sistema? **No**

- Disponibilidad

1. ¿Al solicitar datos al sistema, ya sea de estudiantes o de académicos, estos siempre se encuentran disponibles? **Si**
2. ¿Al solicitar datos al sistema, ya sea de estudiantes o de académicos, estos siempre se encuentran actualizados? **Si**
3. ¿Los datos solicitados al sistema, se presentan de manera clara? **Si**

Luego de aplicar las encuestas los resultados obtenidos fueron los siguientes:

- **Confidencialidad:** En cuanto a la confidencialidad como un todo, los resultados obtenidos como se pueden apreciar en el Gráfico 4-9 indican que esta falla en un 60% aproximadamente, *siendo la principal causa la desinformación acerca de las políticas de los datos por parte de los usuarios y poca existencia de acuerdos de confidencialidad para las personas que acceden a los datos, estos problemas son considerados de gravedad, ya que los usuarios en su mayoría trabajan con datos personales y confidenciales de los alumnos o académicos y en este momento, en algunos casos, no existe ninguna restricción de divulgación de estos.*

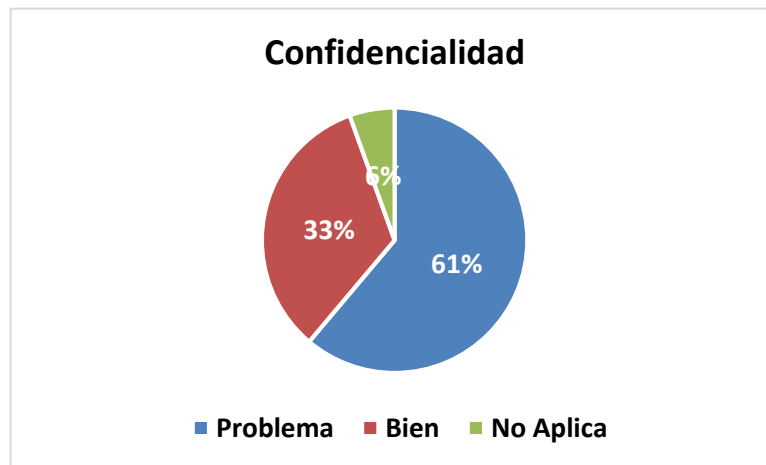


Gráfico 4-9: Resultado en cuanto a confidencialidad de los datos (Fuente: Propia).

- **Integridad:** En cuanto a la integridad como un todo, los resultados obtenidos como se pueden apreciar en el Gráfico 4-10 indican que esta falla en un 30% aproximadamente, *siendo el principal problema la modificación de datos de una persona sin una solicitud previa de este.*

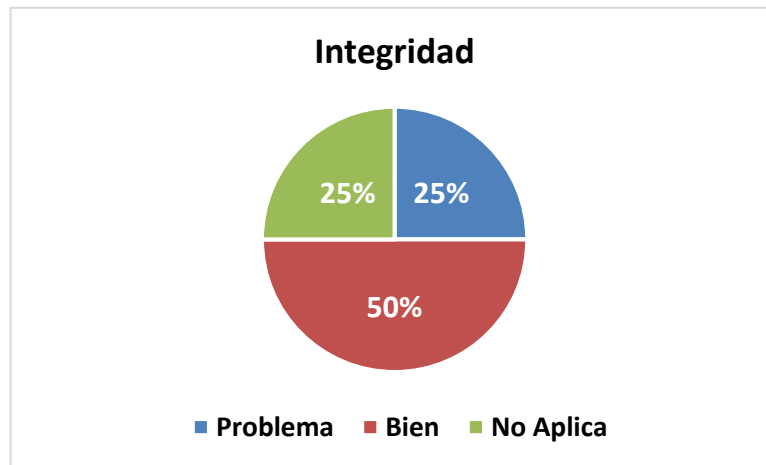


Gráfico 4-10: Resultado en cuanto a integridad de los datos (Fuente: Propia).

- Disponibilidad: En cuanto a la disponibilidad como un todo, los resultados obtenidos como se pueden apreciar en el Gráfico 4-11 indican que esta falla en un 47% aproximadamente, *siendo la principal causa de esto que algunos usuarios consideran que los datos no se presentan de forma clara en el sistema y que estos no siempre se encuentran disponibles ni actualizados.*

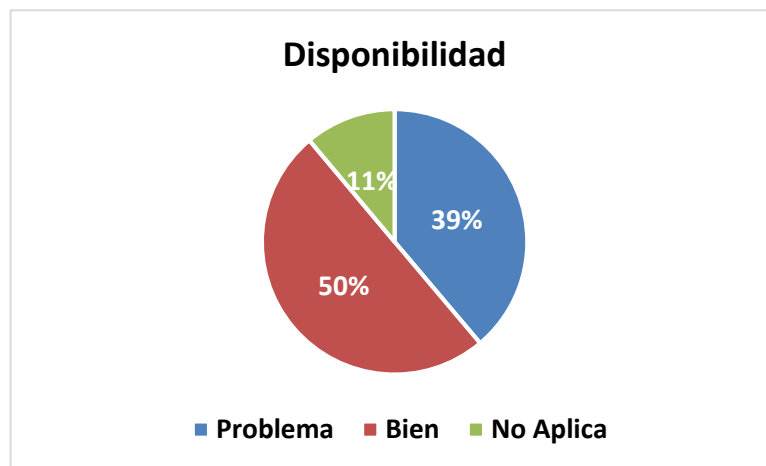


Gráfico 4-11: Resultado en cuanto a disponibilidad de los datos (Fuente: Propia)

Los resultados, como pueden apreciarse en el Gráfico 4-12, presentan que las dimensiones de confidencialidad y disponibilidad se encuentran en niveles de insuficiencia, esto puede deberse a que, en confidencialidad los usuarios tienen poco conocimiento acerca de las

políticas de seguridad de los datos, además de la casi nula existencia de acuerdos de confidencialidad en los usuarios con acceso al sistema. Por otro lado, en la dimensión de disponibilidad los usuarios consideraron que los datos no se presentan de forma clara en el sistema y no siempre se encuentran actualizados.

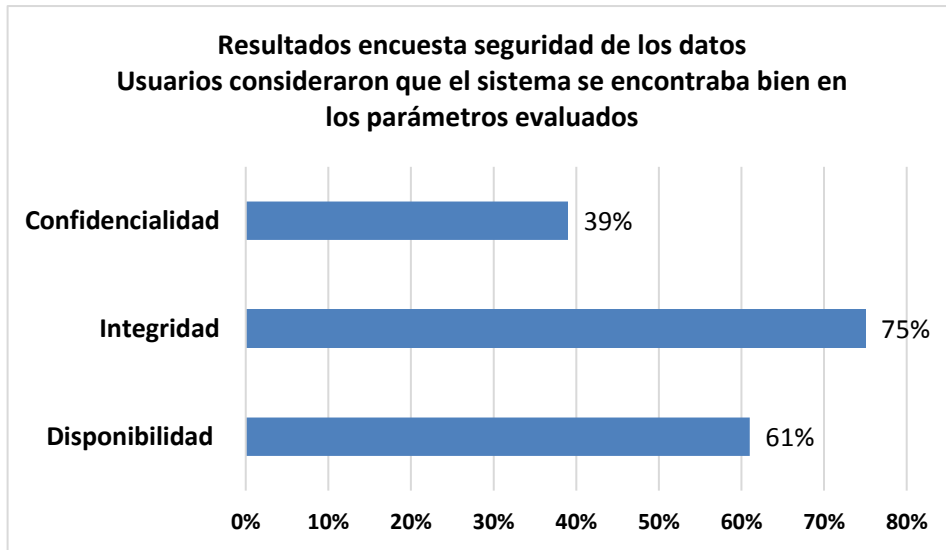


Gráfico 4-12: Resultado encuesta seguridad de los datos (Fuente: propia).

4.5 Revisión del sistema en cuanto a seguridad de los datos

Se realizó una revisión de la base de datos del sistema, inspeccionando 3 usuarios con distintos roles, para esto, como se muestra en la Tabla 4-3, se realizó un seguimiento de los posibles datos que son ingresados, consultados, eliminados o modificados en las distintas tablas, con el fin de encontrar problemas de seguridad de los datos del sistema SIMBAD los cuales no son medibles a través de encuestas, además se pudo verificar algunas de las respuestas entregadas por los usuarios.

Con el fin de medir la consistencia de la base de datos del sistema es que se procedió a inspeccionar distintos roles del sistema SIMBAD, se realizaron distintas pruebas con el fin de medir las dimensiones propuestas en el capítulo anterior (confidencialidad e integridad). Cada tabla fue analizada con 3 roles distintos: el de secretarias, jefes de carrera y el de administrativos.

Se analizaron distintos aspectos, de los cuales se calculó un porcentaje de error de la siguiente forma (Covella, 2005):

$$R = \frac{TPE}{TP} * 100\%$$

Donde:

TPE: Total pruebas con error.

TP: Total de pruebas.

Para evaluar los resultados obtenidos de la revisión se utilizó una rúbrica de 3 posibles criterios:

Excelente → Errores < 3%

Aceptable → 2 % < Errores < 29%

Insuficiente → Errores > 28%

Tabla 4-3: Revisión a la base de datos de SIMBAD (Fuente: Propia).

SEGURIDAD DE LOS DATOS			
Dimensión	Tipo de errores	Roles analizados	Errores encontrados
Confidencialidad	Acceso a datos innecesarios	3	0
	Sesiones mal restringidas	3	0
Integridad	Modificación de datos libremente	3	0
	Eliminación de datos libremente	3	0

Los resultados obtenidos, como se pueden apreciar en el Gráfico 4-13, muestran que, en los aspectos medidos en cuanto a seguridad de los datos, estos se encuentran en parámetros de excelencia, cabe destacar que, de los usuarios revisados, ninguno pudo efectuar acciones que no le correspondiesen, ni realizar modificaciones libremente.

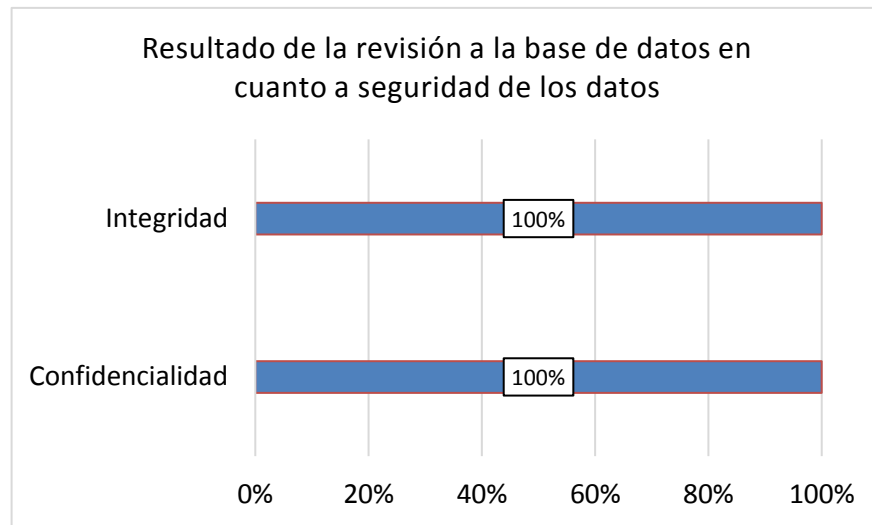


Gráfico 4-13: Resultados revisión base de datos de SIMBAD (Fuente: Propia)

4.6 Conclusión y discusión del capítulo

A modo de conclusión en este capítulo y luego del análisis realizado al sistema, se pudo concluir que el sistema SIMBAD se encuentra muy bien diseñado pues el modelo de la base de datos, los tipos de datos y los formatos de los datos son lógicos y bien estructurados. El sistema está bien diseñado desde el punto de vista lógico, pareciera ser por tanto que el principal problema fue la implementación de la interfaz, siendo poco intuitiva, lo que genera confusiones y errores en los usuarios.

En cuanto a la calidad de los datos los mayores problemas encontrados fueron la estandarización de los datos, ya que estos al provenir de distintas fuentes o ser digitados por diversas personas, no siempre tienen el mismo formato y en este caso el sistema no es capaz de unificar todo. Por otro lado, en algunos casos los datos no están validados, y cuando los datos son registrados de forma manual cabe la posibilidad de que se cometan errores de tipeo.

Los registros vacíos en la base de datos del sistema son un problema también porque de no ser ocupados en su gran mayoría indica que no son funcionales para el sistema y sólo se encuentran ocupando memoria y haciendo más engorroso el proceso.

En cuanto a la seguridad de los datos, los resultados obtenidos por los usuarios muestran la falta de información acerca de las normativas que rigen los datos en la UCSC y a su vez estos, en su mayoría, no poseen ningún acuerdo de confidencialidad de los datos lo que facilita la divulgación o mal uso de los datos registrados en el sistema SIMBAD, ya que de los 6 encuestados, 5 contestaron que no firmaron ningún acuerdo de confidencialidad de los datos, entregando poca seguridad a los alumnos y académicos.

En el siguiente capítulo se describen una serie de procedimientos que van en ayuda de cumplir criterios de calidad y seguridad de los datos, con el fin de mejorar los aspectos negativos encontrados en el diagnóstico realizado.

Capítulo 5 . Procedimientos para la calidad y seguridad de los datos

En el presente capítulo se detallan una serie de procedimientos centrados en los datos activos para el sistema SIMBAD, los cuales buscan mejorar los aspectos negativos encontrados en el capítulo 4 y recalcar los aspectos positivos en cuanto a calidad y seguridad de los datos del sistema, enfocándose en cumplir las dimensiones mencionadas en el capítulo 4 en las secciones 4.1.2 y 4.2.1.

Para poder definir ciertos procedimientos que van en ayuda a la calidad y seguridad del sistema, fue necesario distinguir y tener claro cuáles son los datos activos para el sistema SIMBAD, ya que sin la existencia de estos el sistema no se justificaría y gran parte de sus funcionalidades no funcionarían; para esto en primera instancia se realizó un estudio de las 18 tablas principales de la base de datos que interactúa en el sistema y a su vez se clarificó los datos activos para cada una de estas tablas, revisar Anexo D. La base de datos de sistema está desarrollada en Oracle el cual trabaja con lenguaje PL/SQL.

5.1 Procedimientos de calidad de los datos

Para poder lograr criterios de excelencia en calidad de los datos, mencionados en el capítulo 4 sección 4.1.2, es indispensable seguir una serie de procedimientos en cuanto al uso de los datos. A continuación, se describe una lista de acciones que van en ayuda al cumplimiento de cada una de las dimensiones mencionadas en el capítulo 4 en la sección 4.1.1, ya sea a nivel de sistema como al de usuario, con esto se espera ayudar a mejorar el nivel de calidad de los datos.

5.1.1 Completitud

Como se mencionó anteriormente, completitud busca que los datos estén completos, pero a la vez sean válidos. Se buscó mejorar el nivel de completitud del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez

más a parámetros de excelencia, para ir en ayuda de esto se definieron procedimientos para mejorar los siguientes aspectos:

1. Inexistencia de datos activos vacíos en el sistema

Se espera que el sistema se encuentre lo más completo posible en lo que respecta a los datos, enfocándose principalmente en los primarios ya que estos son los que justifican el funcionamiento del sistema.

I. Procedimientos asociados al sistema:

- Todos los datos activos deberán estar restringidos, en la base de datos del sistema, por la propiedad NOT NULL para así asegurar que no se presenten campos vacíos importantes para el funcionamiento del sistema.
- En el registro de personas nuevas al sistema, para el campo “PERS_E_MAIL” correspondiente al email, el sistema asignará un correo electrónico institucional, evitando campos vacíos, este será la concatenación del primer carácter del primer (PERS_NOM1) y segundo nombre (PERS_NOM2) más el apellido paterno completo (PERS_AP_PATERNO), asignándole el dominio correspondiente a la carrera o cargo asignado en la universidad. En el caso de existir alcance de nombres y dominios; y con el fin de evitar duplicidad de correos se procederá a la concatenación del primer carácter del primer (PERS_NOM1) y segundo nombre (PERS_NOM2) más el apellido paterno completo (PERS_AP_PATERNO) más el primer carácter del apellido materno (PERS_AP_MATERNO), a modo de ejemplificación se presenta la Figura 5-1.

II. Procedimientos asociados a la interfaz usuario:

- El sistema no permitirá cargar los datos si es que se presentan campos vacíos de datos activos, en tal caso el sistema desplegará un mensaje al usuario indicándole “Falta rellenar campos” y destacando estos en color rojo.
- En el registro de personas nuevas al sistema, en el campo “DIRE_DIR1” correspondiente a la dirección principal, en el caso de que la persona viva en un lugar

rural el cual no posee una dirección física específica, será obligación del usuario solicitar una dirección exacta de algún familiar cercano donde se le pueda enviar correspondencia e ingresar en “DIRE_DIR2”, una dirección secundaria la cual será una aproximación de su domicilio rural.

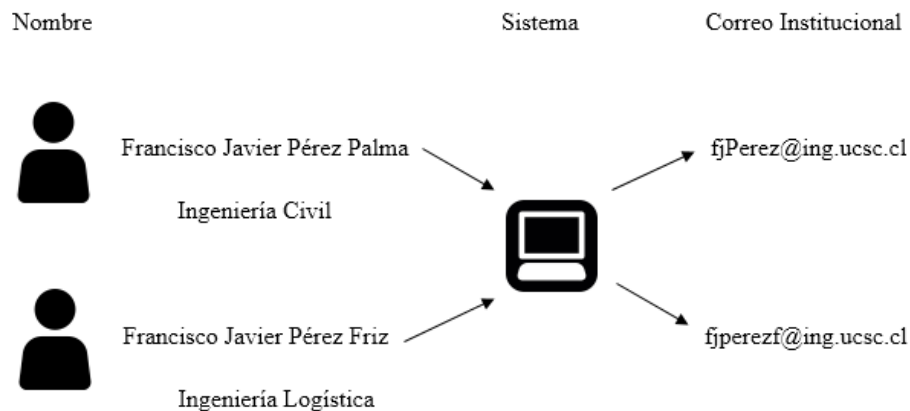


Figura 5-1: Proceso de asignación de correo institucional (Fuente: Propia).

2. Inexistencia de eliminaciones de datos primarios accidentalmente en el sistema

Se espera que no se realicen eliminaciones erróneas de datos activos del sistema, tanto a nivel de base de datos como de usuario, ya que estos pueden generar incompletitudes en el sistema.

I. Procedimientos asociados al sistema:

- Los usuarios que posean rol distinto a de SUPER_USUARIO, R_JEFE_DOCENCIA o R_JEFE_TIT_Y_GRD no tendrán permisos asociados a DELETE o UPDATE de los datos del sistema.
- Los usuarios con rol R_JEFE_DOCENCIA o R_JEFE_TIT_Y_GRD serán los únicos que tendrán permisos totales para la administración de los datos del sistema.
- Los usuarios con rol SUPER_USUARIO serán los únicos que tendrá permisos totales para la administración del sistema.

II. Procedimientos asociados a la interfaz del usuario:

- Los usuarios del sistema ya sean las secretarías académicas, jefes de carrera o administrativos no podrán realizar ningún tipo de eliminación inmediata, exceptuando el caso de Reserva de Salas, para esto el sistema le entregará una sección de solicitud de modificación de datos, donde podrán ingresar la solicitud pertinente la cual será revisada y validada posteriormente por la DARA, esto con el fin de hacer el proceso lo más transparente posible.
- Todos los usuarios del sistema podrán realizar reservas de salas y podrán realizar la eliminación o modificación de las reservas que ellos mismos hayan realizado.

3. Inexistencia de datos innecesarios en el sistema
--

Se espera que el sistema sea lo más preciso y completo posible, para ello es de suma importancia que no existan datos innecesarios, el fin de estos procedimientos es tener lo más optimizado posible los procesos de registros.

I. Procedimientos asociados al sistema:

- Si existen tablas en la base de datos del sistema que posean columnas que en su mayoría no posean datos y que estos no sean considerados datos activos para el sistema, se procederá a la eliminación de éstas con el fin de que no se produzcan confusiones y que los registros se encuentren lo más completos posible. Para determinar que columnas podrán ser eliminadas se recomienda utilizar la siguiente métrica:

Métrica: Eliminación de campos innecesarios (Fuente: propia).

$$M = \frac{\text{Número de registros vacíos}}{\text{Número de registros}} \times 100\%$$

Donde:

Número de registros: total de registros en la tabla.

Número de registros vacíos: total de registros que no poseen datos en la columna en cuestión.

Donde:

Si $M < 60$ y $dato \neq activos \rightarrow$ La columna será eliminada

- Se eliminarán de la tabla DIRECCIONES los campos de DIRE_C_CMNA, DIRE_C_PAIS, ya que estos pueden ser obtenidos con el dato DIRE_C_CIUADAD, creando una relación entre las tablas C_CIUADADES, C_COMUNAS, C_PAISES.

II. Procedimientos asociados a la interfaz del usuario:

- Al realizar un registro de dirección en el sistema, el usuario tendrá solo que digitar la comuna de procedencia y el sistema automáticamente rellenará los campos de ciudad y región, con el fin de agilizar el proceso.

4. La longitud de los campos en el sistema deberá estar debidamente restringida

Se espera que tanto en la base de datos como en la interfaz del sistema no se puedan registrar datos con longitudes distintas a las permitidas por el parámetro, con el fin de minimizar al máximo los errores y que los datos sean lo más completos posible.

I. Procedimientos asociados al sistema:

- En la base de datos del sistema, la tabla PERSONAS en el campo “PERS_RUT”, correspondiente al Rut, deberá estar restringido con una longitud mínima de 7 caracteres numéricos y como máximo 8 caracteres numéricos.
- En la base de datos del sistema, la tabla PERSONAS en el campo “PERS_NOM2”, correspondiente al segundo nombre de la persona, deberá tener una longitud de 60 caracteres, ya que en ocasiones las personas tienen más de 2 nombres.
- En la base de datos del sistema, la tabla DIRECCIONES en el campo “DIRE_TELEFONO_CELULAR”, correspondiente al teléfono celular, deberá tener una longitud única de 10 caracteres numéricos.

II. Procedimientos asociados a la interfaz del usuario:

- En el proceso de registro de personas en el sistema, en el campo de número celular no está permitido ingresar más de 10 números, y en el caso de que este ingrese una cantidad menor de dígitos, el sistema desplegará un mensaje indicando “Error Formato de Teléfono Incorrecto”.
- En el proceso de registro de personas en el sistema, en el campo RUT no permitirá al usuario ingresar más de 8 dígitos y en el caso de que este ingrese una cantidad menor a 7 dígitos, el sistema le desplegara un mensaje indicando “Error faltan números”.
- El sistema deberá indicarle al usuario al momento de realizar el registro de una nueva persona, que en el campo Rut, este debe ser ingresado sin puntos ni dígito verificador.

5.1.2 Consistencia

Como se mencionó anteriormente la consistencia busca la ausencia de diferencias entre comparaciones de representaciones distintas con los mismos datos. Se espera mejorar el nivel de consistencia del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia. Para ir en ayuda de esto se definieron los siguientes procedimientos:

1. Obtención de los datos para el sistema
--

Se espera que todos los datos que se registran en el sistema, sean de la fuente que vengan, siempre sean consistentes y lo más precisos y confiables posible.

I. Procedimientos asociados al sistema:

- Los datos que se registrarán en el sistema de los alumnos que ingresan por proceso de admisión regular a la universidad serán obtenidos del DEMRE, los cuales serán posteriormente confirmados por los mismos alumnos al momento de la matrícula.
- Los datos que se registrarán en el sistema de los alumnos que ingresan por proceso de admisión especial serán completados por los alumnos vía intranet, además se

solicitará adjuntar un certificado de nacimiento. Los datos ingresados serán posteriormente validados por los puntos descritos en la sección 5.1.5.

- Los datos que alimentarán al sistema con los registros de los académicos de la universidad serán obtenidos directamente de la Dirección de Gestión de Personal.

2. Actualización de los datos

Se espera que los datos del sistema siempre se encuentren lo más actualizados posibles, para que con esto en todo momento sean verdaderos y consistentes con la realidad.

I. Procedimientos asociados al sistema:

- El sistema le solicitará cada 6 meses a las personas registradas en el sistema que confirmen sus datos personales más relevantes, estos son: “DIRE_C_CIUDAD” ciudad donde habita, “PERS_E_MAIL” correo electrónico, “DIRE_DIR1” dirección domiciliaria, “DIRE_TELEFONO_CELULAR” número de celular y “DIRE_FONOS” número de red fija.
- El sistema actualizará los datos registrados en la dirección de Gestión de personal semanalmente con el fin de mantener al día los registros del personal.

II. Procedimientos asociados al usuario:

- Para la actualización de los datos del personal como los grados obtenidos o especialidades realizadas, estos deberán acercarse directamente a la dirección de Gestión de Personal y una vez actualizados, el sistema actualizará sus datos en el periodo máximo de 7 días.
- Las evaluaciones realizadas por los docentes deberán ser registradas en el sistema en un periodo máximo de 15 posterior a la fecha de evaluación, con el fin de mantener el sistema lo más actualizado posible y resguardar la información de los alumnos.
- Si es necesario la actualización de los datos de un alumno, este deberá acercarse a la Dirección de Admisión y Registro Académico y rellenar una solicitud en papel la cual

será posteriormente validada y registrada, la actualización del sistema deberá estar realizada en un periodo máximo de 3 días hábiles posteriores a la solicitud.

3. Integración de todos los datos independiente de la fuente

Los datos digitados en el sistema y los datos obtenidos de distintas fuentes serán integrados en el sistema, cumpliendo parámetros previamente acordados, con el fin de evitarán redundancias, problemas de inconsistencias y duplicación de los datos.

I. Procedimientos asociados al sistema:

- El sistema no registrará datos que provengan de otras fuentes que daten de antes del año 1991 ya que este es el año de creación de la institución.
- El sistema al momento de nutrirse de otras fuentes validará que los registros no se encuentren en el sistema con el fin de evitar duplicidad, y en el caso de existir el registro se procederá a la actualización de los datos de ser necesario.
- El sistema sólo se nutrirá de fuentes primarias, tales como: Dirección de Gestión de Personal, para datos de los académicos; profesores a cargo de cursos, para el ingreso de notas al sistema; alumnos, para datos personales; y DEMRE, para datos necesarios para estadísticas; esto con el fin de evitar inconsistencias.
- En el caso de ser necesario recaudar datos de otras fuentes, esta deberá ser validada por la persona o institución en cuestión.

5.1.3 Exactitud

Como se mencionó anteriormente, se buscó mejorar el nivel de exactitud del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia, para ir en ayuda de esto se definieron los siguientes procedimientos:

1. Formato de los datos

Todos los datos del sistema deberán estar almacenados en la base de datos en un mismo formato y a la vez estos deberán ser presentados en el sistema en un mismo e intuitivo formato para el usuario.

I. Procedimientos asociados al sistema:

- Los datos serán registrados en mayúscula, si estos son escritos en minúscula el sistema los transformará unificando todo en un mismo formato.
- En el sistema solo se permitirán números naturales.
- El formato para los registros de tipo fecha será el de dd/mm/aaaa, si se ingresan registros con otro formato, este deberá ser capaz de transformarlos unificando todo en un mismo formato.
- El formato para los registros de tipo hora será el de hh/mm/ss, si se ingresan registros con otro formato, este deberá ser capaz de transformarlos unificando todo en un mismo formato.

II. Procedimientos asociados a la interfaz del usuario:

- El formato con el que se desplegarán los datos no numéricos en el sistema será fuente Arial mayúscula tamaño 12px, si estos son escritos en minúscula el sistema los transformará unificando todo en un mismo formato.
- El formato con el que se desplegarán los datos numéricos en el sistema será fuente Arial tamaño 12px.
- El sistema deberá indicar al usuario, de manera clara, los formatos en los que el sistema registrará los datos, ya sean las horas (hh/mm/ss), las fechas (dd/mm/aaaa)

2. Los datos deben ser estandarizados en un formato único

El sistema unificará y estandarizará los datos ingresados al sistema, sin importar la fuente de la que provengan, ya sean estos digitalizados o extraídos de distintas bases de datos.

I. Procedimientos asociados al sistema:

- Todos los registros que se realicen en el sistema se almacenarán en la base de datos de este, y todos los datos serán guardados en un mismo formato previamente descrito en el punto 1 de esta sección.
- Todos los registros provenientes de otros sistemas, serán almacenados en la base de datos del sistema y estos deberán unificarse con los demás datos, adecuándose a los formatos establecidos. A modo de ejemplo se presenta la Figura 5-2.

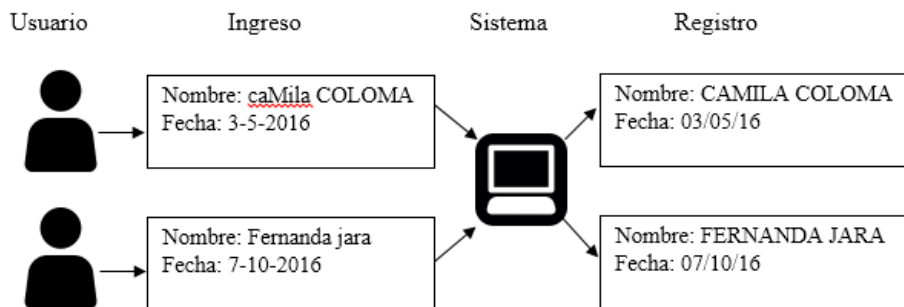


Figura 5-2: Ejemplificación de la estandarización de los datos (Fuente: Propia).

3. Los rangos de los parámetros en el sistema deberán estar correctamente limitados

Los parámetros en el sistema deberán estar correctamente limitados en cuanto a los rangos con el fin de evitar equivocaciones.

I. Procedimientos asociados al sistema:

- En la tabla PERSONAS si el nuevo registro es de tipo (PERS_SUB_TIPO) “ALUM”, se deberá restringir adecuadamente la fecha de nacimiento (PERS_F_NACIMIENTO), para lo cual se aconseja la utilización de la siguiente métrica:

Métrica: Año mínimo de ingreso para alumnos.

$$Edad = A_{AC} - A_{NAC}$$

Donde:

$$A_{AC} = \text{Año actual}$$

$$A_{NAC} = \text{Año de nacimiento}$$

Donde:

Si $Edad < 17 \rightarrow$ El registro no es permitido

Si $Edad \geq 17 \rightarrow$ El registro es permitido

- En la tabla PERSONAS si el nuevo registro es distinto de tipo (PERS_SUB_TIPO) "ALUM" se deberá restringir adecuadamente la fecha de nacimiento (PERS_F_NACIMIENTO), para lo cual se aconseja la utilización de la siguiente métrica.

$$Edad = A_{AC} - A_{NAC}$$

Donde:

$$A_{AC} = \text{Año actual}$$

$$A_{NAC} = \text{Año de nacimiento}$$

Donde:

Si $Edad < 18 \rightarrow$ El registro no es permitido

Si $Edad \geq 18 \rightarrow$ El registro es permitido

II. Procedimientos asociados a la interfaz del usuario:

- Si el usuario al momento de registrar a un alumno en el sistema ingresa un año de nacimiento menor al establecido por la base de datos, se desplegará un mensaje indicando "Año de nacimiento inválido"

- Si el usuario al momento de registrar a un funcionario en el sistema ingresa un año de nacimiento menor al establecido por la base de datos, se desplegará un mensaje indicando “Año de nacimiento inválido”

4. Clarificación de los campos de llenado en el sistema

Los campos de llenado en el sistema deberán estar claramente descritos indicando la descripción del dato con el formato de este.

I. Procedimientos asociados a la interfaz del usuario:

- En el módulo Mantenedor de Personas en el sistema, en la sección de personas en el campo “Sub-tipo” se deberá detallar lo que significa cada una de las opciones disponibles (ALPR, ALUM, PROF, POST, OTRO, AUTO, FUNC).
- En el módulo Mantenedor de Personas en el sistema, en la sección de personas en el campo “Estado” se deberá dar la opción de revisar una lista con las opciones disponibles.
- En el módulo de Mantenedor de Personas en el sistema, en la sección de antecedentes en el campo de “fecha de nacimiento” se deberá indicar que el formato correcto es DD-MM-AAAA.
- En el módulo Mantenedor de Personas en el sistema, en la sección de antecedentes en el campo de “Sexo”, se deberá indicar que “F” se refiere a femenino y “M” a masculino.
- En el módulo Mantenedor de Personas en el sistema, en la sección de antecedentes en el campo “Nacionalidad”, “Estado Civil”, “Estab. Ed. Media” y “Tipo de Ingreso” se deberá dar la opción de revisar una lista con las opciones disponibles.
- En el módulo Mantención de Salas del sistema en “Categorías” se deberá dar la opción de revisar una lista con las opciones disponibles.
- En el módulo Mantención de Cursos en el sistema en los campos “Tipo de curso” y “Sede” se deberá dar la opción de revisar una lista con las opciones disponibles. A modo de ejemplo se presenta la Figura 5-3.

Sexo:

(F: FEMENINO/ M: MASCULINO)

Figura 5-3: Ejemplificación de la descripción de los datos en el campo sexo (Fuente: Propia).

5.1.4 Unicidad

Como se mencionó anteriormente, se buscó mejorar el nivel de unicidad del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia. Para ir en ayuda de esto se definieron los siguientes procedimientos:

<p>1. Inexistencia de duplicidad en los datos</p>
--

El sistema deberá estar completamente libre de duplicidades para así evitar que se puedan generar contradicciones en los datos.

I. Procedimientos asociados al sistema:

- El diseño de la base de datos deberá estar correctamente normalizado, cumpliendo con las 3 formas normales:
 1. Primera forma normal: Todos los atributos de las tablas deberán ser atómicos, esto quiere decir que los elementos del dominio son indivisibles. Las tablas deberán tener una única clave primaria, los datos de la tabla que no sean clave deberán tener dependencia funcional.
 2. Segunda forma normal: Crear tablas independientes para conjunto de datos que se apliquen a varios registros y relacionarlas con una clave externa.
 3. Tercera forma normal: Eliminar todos los campos que no dependan directamente de alguna clave.
- En la tabla Personas de la base de datos del sistema, se deberá corregir el campo de “PERS_SUB_TIPO”, que indica el tipo de función que desempeña en la universidad

una persona (alumno, docente, secretaria, entre otros), ya que en el caso actual solo permite ingresar una opción por persona y puede darse el caso de personas que cumplan más de una función.

- En la base de datos del sistema el “Rut” de la persona será clave primaria para todos los datos asociados a ella, ya sean antecedentes, direcciones, cursos, notas, entre otros.

II. Procedimientos asociados a la interfaz del usuario:

- Si al momento de realizar algún registro en el sistema el Rut a ingresar ya existe, se desplegará un mensaje indicando “Error, Rut ya existente”.

2. Garantizar correctitud de las claves primaria en la base de datos

Para asegurar que no se produzca duplicidad de los datos y no se generen contradicciones, es primordial tener bien definidas las claves primarias de las principales tablas en la base de datos del sistema.

I. Procedimientos asociados al sistema:

- Para la tabla CARRERAS de la base de datos del sistema, el parámetro “CARR_CODIGO” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla PLANES_CARR de la base de datos del sistema, el parámetro “PLAN_CODIGO” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla CURSOS de la base de datos del sistema, el parámetro “CURS_CODIGO” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla RESTRIC_INSC de la base de datos del sistema, los parámetros “RI_CURS_CODIGO” y “RI_C_CARRERA” deberán cumplir el papel de claves primarias para la tabla.
- Para la tabla PROFESORES_CURSO de la base de datos del sistema, los parámetros “P_C_C_CURSO” y “P_C_PROF_RUT” deberán cumplir el papel de claves primarias para la tabla.

- Para la tabla PERSONAS de la base de datos del sistema, el parámetro “PERS_RUT” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla DIRECCIONES de la base de datos del sistema el parámetro “DIRE_PERS_RUT” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla RAMOS de la base de datos del sistema el parámetro “RAMO_C_RAMO” deberá cumplir el papel de clave primaria para la tabla.
- Para la tabla SALAS de la base de datos del sistema el parámetro “SALA_C_SALA” deberá cumplir el papel de clave primaria para la tabla.

5.1.5 Validez

Como se mencionó anteriormente, se buscó mejorar el nivel de validez del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia, para ir en ayuda de esto se definieron los siguientes procedimientos:

1. Asegurar que los datos ingresados son reales
--

Se espera que los datos ingresados en el sistema sean verdaderos y para esto es necesario que sean validados en algunas ocasiones.

I. Procedimientos asociados al sistema:

- Las direcciones principales, ya sean domiciliarias o de algún tipo de institución, serán validadas antes de ser registradas en el sistema, por medio de Google Maps Geocoding API con el fin de asegurarse que la dirección exista.
- Los correos electrónicos no institucionales deberán ser validados antes de ser registrados en la base de datos del sistema, por medio de algún sistema web que proporcione el servicio, algunos de los sistemas propuestos son:
 - <http://www.email-unlimited.com/tools/verify-email.aspx>
 - <http://www.comprobarcorreo.com/comprobarcorreo-en-tu-web>
 - <http://www.verifyemailaddress.org>

- Los Rut ingresados al sistema deberán ser validados en la base de datos por medio de la función de PL/Sql:

```

FUNCTION VALIDA_RUT ( p_rut in number, p_dv in varchar2) RETURN boolean IS
    v_rut varchar2(9);
    v_dv1 varchar2(1);
    v_dv2 varchar2(1);
BEGIN
    v_rut := lpad(p_rut,9,'0');
    v_dv1 := p_dv;
    v_dv2 := substr('123456789K0',11-mod(to_number(substr(v_rut,1,1))*4 +
        to_number(substr(v_rut,2,1))*3 +
        to_number(substr(v_rut,3,1))*2 +
        to_number(substr(v_rut,4,1))*7 +
        to_number(substr(v_rut,5,1))*6 +
        to_number(substr(v_rut,6,1))*5 +
        to_number(substr(v_rut,7,1))*4 +
        to_number(substr(v_rut,8,1))*3 +
        to_number(substr(v_rut,9,1))*2,11),1);
    if v_dv1 = v_dv2 then
        return(TRUE);
    else
        return(FALSE);
    end if;
    exception
        when others then
            return(FALSE);
END;
```

II. Procesos asociados a la interfaz del sistema:

- Si la dirección ingresada en el sistema no es posible de verificar, aparecerá un mensaje indicando “Dirección no encontrada”.

- Si el correo electrónico ingresado no es validado, aparecerá un mensaje indicando “Correo electrónico no existe”.
- Si el rut ingresado en el sistema no es válido, se desplegará un mensaje indicando “Rut inválido”.

III. Procesos asociados al usuario:

- Al momento de que un usuario tenga que registrar a una persona en el sistema será estrictamente necesario que se le solicite el carnet de identidad, este deberá estar actualizado y en condiciones óptimas, de lo contrario no se podrá realizar el ingreso.
- Al momento de registrar a un usuario al sistema, el nombre deberá ser validado por medio de un certificado de nacimiento obtenido por medio del registro civil.

2. Asegurar que los datos digitados en el sistema no poseen errores de tipeo

Se espera que se pueda reducir al máximo los errores que van asociados a una mala digitación.

I. Procesos asociados al sistema:

- Se utilizarán las funciones de geo-autocompletar de la API de Google Maps, para ayudar al usuario al momento de ingresar direcciones.
- Para evitar errores básicos de tipeo se corroborará que el correo electrónico cumpla con un modelo típico de email, para esto se realizará una validación en la base de datos por medio de la siguiente línea de código:

```
SELECT      regexp_substr('correo@dominio.com','[a-zA-Z0-9._%~]+@[a-zA-Z0-9._%~]+\.[a-zA-Z]{2,4}')
FROM DUAL
```

El retorno de la ejecución de la sentencia Select, retornara la dirección de correo electrónico si es válida, de lo contrario retornará NULL. Cabe destacar que esta línea de código es para bases de datos desarrollada en Oracle.

II. Procesos asociados a la interfaz del usuario:

- Si el usuario digita una dirección en el sistema, se le sugerirán ubicaciones en tiempo real a medida que escribe, gracias a la función de geo-autocompletar.
- Si el usuario digita un correo electrónico que no cumpla con el formato mínimo, el sistema desplegará un mensaje indicando “Correo electrónico mal escrito”.

5.2 Procedimientos de seguridad de los datos

Para poder lograr seguridad de los datos es indispensable seguir una serie de procedimientos en cuanto al uso de estos, para ello a continuación se describe una lista de acciones que van en ayuda al cumplimiento de cada una de las dimensiones mencionadas en el capítulo 4 sección 4.2.1, ya sea a nivel de sistema como al de usuarios, con esto se espera ayudar a lograr alcanzar criterios de evaluación de excelencia mencionado en el capítulo 4 sección 4.2.2.

5.2.1 Confidencialidad

Como se mencionó anteriormente confidencialidad es la propiedad que busca prevenir la divulgación de información a personas o sistemas no autorizados. Se espera mejorar el nivel de consistencia del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia. Para ir en ayuda de esto se definieron los siguientes procedimientos:

I. Procedimientos asociados a los usuarios:

- Los usuarios que tengan acceso al sistema deberán firmar un acuerdo de confidencialidad acerca de todos los datos que se encuentren registrados en el sistema.
- Cada usuario al momento de registrarse en el sistema SIMBAD, deberá digitar personalmente una clave que como mínimo contenga 4 caracteres y 2 dígitos para

poder acceder al sistema, con el fin de responsabilizar a cada persona por los movimientos realizados en su sesión.

- La DARA deberá dar a conocer a todos los usuarios que trabajen con los módulos que administra el sistema, las políticas o normativas que la universidad tiene en cuanto a la seguridad de los datos.

II. Procedimientos asociados al sistema:

- Se bloqueará la opción de impresión de pantalla “Imp Pan”, con el fin de impedir la divulgación o el mal manejo de información confidencial.
- No podrán ser impresos ni descargados documentos que contengan información personal de alumnos docentes o administrativos, por usuarios que posean un rol distinto al de R_JEFE_DOCENCIA.
- Los usuarios con rol R_JEFE_DOCENCIA serán los únicos que tengan acceso completo al sistema.
- Los usuarios con roles distintos al de R_JEFE_DOCENCIA deberán estar restringidos, acotándolos exclusivamente a las funciones necesarias para su función en la universidad.

5.2.2 Integridad

Como se mencionó anteriormente integridad es la propiedad que busca mantener a los datos libres de modificaciones no autorizadas. Se espera mejorar el nivel de consistencia del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4, y con esto acercarse cada vez más a parámetros de excelencia, para ir en ayuda de esto se definieron los siguientes procedimientos:

I. Procedimientos asociados a los usuarios:

- Ningún usuario del sistema podrá realizar modificaciones de los datos personales de alguna persona registrada en el sistema sin una previa solicitud del individuo en cuestión.

- Todo cambio realizado en cuanto a datos en el sistema, deberá quedar registrado de forma automática en una bitácora digital, donde solo tendrán acceso, usuarios con rol de SUPER_USUARIO, R_JEFE_DOCENCIA y R_JEFE_TIT_Y_GRD, donde se indicará la fecha, la persona que realizo la modificación y el código de la solicitud que se realizó.

II. Procedimientos asociados al sistema:

- Sólo los usuarios con rol de SUPER_USUARIO, R_JEFE_DOCENCIA o R_JEFE_TIT_Y_GRD podrán modificar o eliminar datos del sistema.
- Si un usuario con rol distinto al de SUPER_USUARIO, R_JEFE_DOCENCIA o R_JEFE_TIT_Y_GRD desea modificar o eliminar algún registro del sistema, deberá enviar una solicitud a la DARA, la cual estará sujeta a aprobación.

5.2.3 Disponibilidad

Como se mencionó anteriormente disponibilidad es la característica, cualidad o condición de los datos de encontrarse a disposición de quienes deben acceder a ella, ya sean personas, procesos o aplicaciones. Se espera mejorar el nivel de consistencia del sistema, enfocándose principalmente en los problemas encontrados en el capítulo 4 y con esto acercarse cada vez más a parámetros de excelencia. Para ir en ayuda de esto se definieron los siguientes procedimientos:

I. Procedimientos asociados al sistema:

- Será indispensable la realización de respaldos de los datos del sistema, con el fin de anteponerse a posibles fallas o pérdida de datos en el sistema. Las políticas de respaldo del sistema se detallarán en el siguiente capítulo.
- Los distintos usuarios del sistema sólo tendrán acceso a los módulos y datos necesarios para la realización de su función en la universidad. Las políticas de acceso al sistema se detallarán en el siguiente capítulo.

II. Procedimientos asociados a usuarios:

- Se deberá entregar una copia del manual de usuario correspondiente a cada usuario del sistema.

- Se deberá realizar una capacitación a usuarios nuevos del sistema, mostrando las funcionalidades del módulo correspondiente.
- De presentar problemas con la disponibilidad de datos en el sistema, se deberán contactar con el encargado adecuado. Las políticas de mantenimiento al sistema se detallarán en el siguiente capítulo.

5.3 Conclusión y discusión del capítulo

A modo de conclusión en este capítulo se puede concluir que a pesar que algunos de los procedimientos mencionados anteriormente hoy en día se encuentran en funcionamiento, si se aplicaran en su totalidad los procesos descritos tanto a los usuarios como al sistema, este podría mejorar los aspectos negativos que fueron encontrados en el capítulo 4, mejorando tanto la calidad como la seguridad de los datos del sistema y a su vez mejorar los niveles de evaluación.

Uno de los principales procedimientos que se deben implementar en el sistema son los de confidencialidad, ya que el sistema almacena datos personales y confidenciales, tanto de alumnos como académicos, los cuales pueden ser usados de manera incorrecta o divulgados sin restricción alguna.

Para disminuir errores de digitación al momento del ingreso de datos al sistema es de suma importancia implementar los procedimientos de exactitud y validez, principalmente los procedimientos que aseguren que los datos ingresados se encuentren en rangos reales y los que van en ayuda de los errores de tipeo como la de autocompletación de campos.

Uno de los problemas encontrados en el capítulo anterior es que la implementación del sistema no es muy intuitiva, por lo que es aconsejable implementar los procedimientos propuestos en exactitud, mejorando la experiencia del usuario con el sistema.

Capítulo 6 . Políticas para el sistema SIMBAD

En el presente capítulo se presentan las distintas políticas del sistema con el fin de mejorar el funcionamiento de este.

6.1 Políticas de acceso

A continuación, se describen las políticas de accesos que tiene el sistema SIMBAD como a su vez los distintos roles que posee el sistema.

6.1.1 Roles del sistema

A continuación, se describen los distintos roles que se encuentran presentes en el sistema.

1. Roles de mantención: Son accesos para la realización de cambios en el sistema. A continuación se detalla en la Tabla 6-1.

Tabla 6-1: Rol de mantención del sistema SIMBAD (Fuente: Propia).

ROLES	DESCRIPCION
SUPER USUARIO	Acceso a todo el sistema.

Los usuarios que tienen acceso a este rol son:

- Administrador de base de datos.
 - Jefe de unidad de desarrollo y soporte de sistemas académicos estudiantiles (DYSSAE).
2. Roles acceso y uso: Son los distintos accesos que tienen los usuarios al sistema, dependiendo del cargo o funciones que realicen, en la Tabla 6-2 se detallan los roles actuales que presenta el sistema.

Tabla 6-2: Roles de uso del sistema SIMBAD (Fuente: Propia).

ROLES DE USO DEL SISTEMA SIMBAD	
ROLES	DESCRIPCION
R_ADMISION	Acceso al módulo de admisión
R_AYUD_DOCENCIA	Acceso al ayudante de Docencia
R_BIENESTAR	Acceso al personal de la DAE
R_COBRANZA	Acceso al Personal de Gestión Financiera estudiantil
R_CONSULTA	Acceso a consulta en las tablas del Sistema
R_COORDAREA	Acceso al coordinador Docente
R_DECANO	Acceso al Decano
R_DIRECTOR_ESCUELA	Acceso al Director de Escuela
R_DOCENCIA	Acceso al personal de Docencia
R_JEFE_DEPT_GESTION_FINANCIERA	Acceso especial para el Jefe de Gestión Financiera
R_JEFE_DOCENCIA	Acceso especial para el Jefe de Docencia
R_JEFE_TIT_Y_GRD	Acceso especial para el Jefe de títulos y grados
R_JEFECARR	Acceso al Jefe de carrera de Pregrado
R_JEFECARR_POST	Acceso al Jefe de carrera de Postgrado
R_MATRICULA	Accesos al módulo de matrícula
R_REGISTRO	Acceso al módulo de registro académico
R_SECRETARIA	Acceso para las secretarias
R_SECRETARIO_ACADEMICO	Acceso al Secretario Académico
R_TITULOS_GRADOS_SECRE	Acceso al módulo de títulos y grados para las secretarias.

6.1.2 Solicitudes de permisos en el sistema

A continuación, se detallan los procedimientos a seguir para solicitar permisos de accesos en el sistema.

1. Usuarios nuevos en el sistema: Para solicitar acceso a usuarios nuevos al sistema, se deberá enviar una solicitud a la Dirección de Docencia y esta una vez evaluada la solicitud enviará la petición a la DSI.
2. Permisos especiales a un usuario: Si un usuario desea acceder a algún módulo al que no tenga permisos, deberá enviar una solicitud a la Dirección de Docencia y si esta evalúa que es necesaria para sus funciones, solicitará a la DSI los permisos de acceso para el usuario.

6.2 Políticas de documentación

A continuación, se detallan los distintos manuales que deberá tener el sistema. Cabe destacar que, de realizarse modificaciones al sistema, todos los manuales mencionados deberán ser actualizados.

- Manual de ingeniería: En este manual deberán estar especificados claramente los requisitos funcionales y no funcionales del sistema, el diseño de este; los pasos a seguir para la instalación en los servidores y en las máquinas en que será utilizado el sistema; la codificación del sistema que detallará la documentación del código y las pruebas al sistema que describirán los procedimientos y resultados de las pruebas realizadas al sistema. Si bien actualmente existe una documentación del sistema SIMBAD, esta no se encuentra completa ni actualizada.
- Manual de usuario: En este manual deberán estar especificados claramente todos los módulos presentes en el sistema, junto a una detallada explicación de su funcionamiento. Actualmente existen manuales de usuario, sin embargo, no se encuentran actualizados ni son muy detallados.
- Manual de usuario por rol: Se deberá tener un manual para cada rol presente en el sistema, donde se especificará claramente el funcionamiento de los módulos que se

pueden acceder, con el fin de entregar una copia correspondiente a cada usuario del sistema. Actualmente no existe un manual para cada rol del sistema.

6.3 Políticas de mantención

Será considerado como mantención del sistema, todas acciones que impliquen modificaciones, correcciones, mejoras o adiciones al sistema, solicitado algún usuario. La mantención al sistema sólo podrá ser realizada por el soporte de aplicaciones de la DSI.

A continuación, se detallan las políticas de mantención para el sistema:

- Problemas en registro de notas: Si se presenta un problema en el registro de notas en el sistema, el usuario que detecte la falla deberá informar inmediatamente a la DARA y esta luego de evaluar el problema, informará al soporte de aplicaciones de la DSI para que realice un seguimiento en el sistema.
- Problemas con el funcionamiento del sistema: Si el usuario detecta problemas en el funcionamiento del sistema deberá comunicarse inmediatamente con el soporte de la facultad a la que pertenezca y este deberá resolver el problema, de no ser así soporte de aplicaciones de la DSI será informada.
- Problemas con la programación del sistema: Si el usuario encuentra que el sistema presenta problemas de inconsistencia en los datos entregados, deberá informar inmediatamente a secretaría de la DARA, quien luego de evaluar la demanda informará a soporte de aplicaciones de la DSI.
- Mantención de servidores: La mantención de servidores donde se aloja el sistema deberá ser realizada fuera de horario académico, este comprenderá de lunes a sábado desde las 21:00 horas hasta las 8:00 horas, de no ser posible esto, se les informará a todos los usuarios, mediante correo electrónico, la hora y fecha en que el sistema se encontrará fuera de uso por motivos de mantención. Se realizará un mantenimiento preventivo general mensual de los servidores donde: se analizará el entorno, el estado físico de la máquina, verificará el estado de conectores y fuentes eléctricas, se comprobará el funcionamiento del servidor, se realizará una limpieza externa e interna del servidor, de ser necesario se realizarán lubricaciones.

Modificación de algún módulo del sistema: De estimar necesario alguna modificación, eliminación o agregación de una o más funcionalidad del sistema, será necesario enviar una solicitud detallada vía correo electrónico al jefe de unidad del DYSSAE. De ser aceptada la solicitud se procederá a la modificación de un sistema de prueba, al cual se le realizarán diversas pruebas, asegurando un buen funcionamiento y el cumplimiento de los requerimientos solicitados. Una vez que las pruebas realizadas sean exitosas, se procederá a poner en funcionamiento el sistema con las distintas modificaciones. Se elaborará o modificará la documentación necesaria.

Para cualquier tipo de mantenimiento será necesario establecer una etapa de prueba, que permita validar la modificación o adición al sistema

6.4 Políticas de respaldo

A continuación, se detallan las políticas de respaldo para el sistema las cuales se realizan por posibles fallos en el sistema.

- Respaldos lógicos: Este tipo de respaldo deberá realizarse diariamente, fuera de horario académico, esto comprende de lunes a sábado desde las 21:00 horas hasta las 8:00 horas, en este sólo se extraen los datos de las tablas de la base de datos del sistema utilizando comandos SQL.
- Respaldos físicos: Este tipo de respaldo deberá realizarse semanalmente y en este se copian los archivos que soportan la base de datos del sistema, se deberá realizar fuera de horario académico, esto comprende de lunes a sábado desde las 21:00 horas hasta las 8:00 horas.
- Respaldos en cinta: Este tipo de respaldo se realiza mensualmente y en este se copian los respaldos en una cinta.

6.5 Conclusión y discusión del capítulo

A modo de conclusión en este capítulo se puede valorizar la importancia que tienen las políticas para cualquier tipo de sistema, ya que gracias a estas se puede reaccionar de manera correcta antes algún tipo de falla o solicitud que se presente en el sistema, en su mayoría el sistema SIMBAD tiene bien definidas sus políticas, sin embargo, documentos importantes no se encuentran actualizados o no existen, como es el caso de la documentación, la cual no se encuentra actualizada y no existe un manual de usuario por rol. Por otro lado, es aconsejable que a cada usuario del sistema se le informe de las políticas en funcionamiento para que así se pueda actuar de la mejor forma ante cualquier imprevisto.

Capítulo 7 Conclusión y recomendaciones

Los planes de DG para las organizaciones son de gran ayuda, ya que estos buscan asegurar la eficacia de los datos, minimizan considerablemente los errores en la toma de decisiones basadas en los datos del sistema. Por otra parte, las realizaciones de estos ayudan a que los usuarios puedan reaccionar de mejor manera en el caso de que ocurra algún inconveniente entregándoles confianza. El asegurar una correcta seguridad de los datos en el sistema ayuda a proteger los datos sensibles de este, dándoles mayor seguridad a las personas que registran sus datos así como mayor compromiso a los usuarios que manipulan o hacen uso de éstos.

De acuerdo a lo propuesto en el Capítulo 1 en lo que respecta a los objetivos específicos se puede concluir lo siguiente:

- Objetivo específico 1: Se estudiaron diversos documentos relacionados con planes de DG, en los cuales se pudo conocer la cantidad de beneficios que tiene la realización de estos para las organizaciones, además de estudiar las distintas metodologías y la estructura que tiene un plan de DG. Se generó una base de conocimiento donde se pudo decidir el área a trabajar (calidad y seguridad de los datos) y la forma de abarcar los distintos puntos para la realización del plan de DG.
- Objetivo específico 2: Luego de la realización de entrevistas a distintos usuarios se pudo decidir trabajar con un sistema completo y de gran ayuda para la UCSC como lo es el sistema SIMBAD. De la realización de las encuestas y de la revisión del sistema se pudo obtener un diagnóstico actual del sistema, el cual estuvo muy cercano a parámetros de excelencia. Los principales problemas encontrados en calidad de los datos fue que el sistema es muy propenso a que se cometan errores de digitación, por lo que se debería tratar de disminuir al máximo posible el riesgo de que esto ocurra. En cuanto a seguridad de los datos el principal problema que se presenta actualmente es la poca confidencialidad que existe con el manejo de éstos.

- Objetivo específico 3: Se definieron una serie de procedimientos, tanto para asegurar la calidad como la seguridad de los datos en el sistema, cabe destacar que SIMBAD cumple con gran parte de los procedimientos propuestos, sin embargo, de llevar a cabo todos los procesos, sus niveles de calidad aumentarían considerablemente.
- Objetivo específico 4: Se definieron políticas para el sistema, tanto de acceso, documentación, mantención y respaldo; se espera que estas sean de gran ayuda tanto para el funcionamiento como para los usuarios del sistema.
- Objetivo específico 5: Se documentó todo el proceso de revisión y determinación de políticas de calidad y seguridad de los datos de SIMBAD.

En aspectos de seguridad y calidad de los datos, SIMBAD se encuentra muy bien catalogado dentro de los parámetros de medición que fueron propuestos para este proyecto, como se mencionó en el Capítulo 4.

Con respecto a la calidad de los datos, uno de los principales problemas que se encontraron es que el sistema es muy propenso a que en el momento de la digitación de datos se comentan errores, esto podría mejorar considerablemente si se considerara la implementación de los procedimientos propuestos en el Capítulo 5, como lo es la limitación de los rangos de los parámetros, la validación de los datos ingresados y las funcionalidades de autocompletación de campos.

Al momento de comparar los resultados obtenidos de las entrevistas realizadas y de la revisión al sistema se pudo concluir que el sistema está bien diseñado, sin embargo, desde el punto de vista lógico y de los datos, la interfaz del sistema es poco intuitiva, lo cual podría mejorar al aplicarse algunos de los procedimientos que se mencionan en el capítulo 5, como la clarificación de los campos y la estandarización de los datos, además de una capacitación respecto del funcionamiento del sistema y la entrega de la documentación correspondiente al módulo asignado.

En cuanto a la seguridad de los datos del sistema, algunos de los principales problemas encontrados fueron la poca presencia de limitaciones legales que tienen los usuarios del sistema para la divulgación y uso incorrecto de datos personales y confidenciales

almacenados en el sistema, esto cambiaría al aplicarse los procedimientos de confidencialidad propuestos en el Capítulo 5, principalmente lo que tiene relación a acuerdos de confidencialidad para cada usuario del sistema SIMBAD, y la limitación de impresiones de los datos.

Como sugerencia, se les debe dar mayor apoyo a los usuarios nuevos en el sistema SIMBAD, con la entrega de un manual de usuario y a su vez una capacitación del funcionamiento y políticas del sistema, como también asegurarse que en ningún caso los datos registrados en el sistema serán divulgados firmando un acuerdo de confidencialidad por parte de todos los usuarios del sistema SIMBAD.

Una de las mayores dificultades que se presentaron en el desarrollo del proyecto, fueron el concordar tiempos de reuniones con terceras personas, y por otra parte tener que esperar respuestas o permisos especiales para poder avanzar en el proyecto.

Bibliografía

- Alvin, T. (2006). *Revolutionary Wealth*. Estados Unidos: Editorial Knopf.
- BBC. (2016, Agosto 21). Retrieved from <http://www.bbc.co.uk/schools/gcsebitesize/ict/databases/6datasecurityrev1.shtml>
- Cai, L., & Zhu, Y. (2015). The Challenges of Data Quality and Data Quality Assessment in the Big Data Era. *Data Science Journal*, 1-10.
- Carnegie Mellon University . (2016, Agosto 17). Retrieved from <http://www.library.cmu.edu/datapub/dms/data/secure>
- Cauca, U. d. (2016, Junio 3). *Facultad de ciencias contables, económicas y administrativas*. Retrieved from <http://fceca.unicauca.edu.co/old/siconceptosbasicos.htm>
- Chiavenato, I. (2006). *Introducción a la teoría general de la administración*. McGraw-Hill Interamericana.
- Copro. (2016, junio 10). Retrieved from http://copro.com.ar/Gestion_de_datos.html
- Da Veiga, A., & Eloff, J. (2010). A framework and Assessment Instrument for Information Security Culture. *Computers & Security*, 196–207.
- Dhillon, G., & Backhouse, J. (2000). Information System Security Management in the New Millennium. *Communication of the ACM*, 125-128.
- Dirección de Servicios Informaticos. (2007). *Modelo Base de Datos*. Concepción.
- Dirección de Servicios Informáticos de la UCSC. (2016, Mayo 30). *DSI*. Retrieved from <http://dsi.ucsc.cl>
- Dyché, J. (2007). *A Data Governance Manifesto: Designing and Deploying Sustainable Data*. Los Angeles.
- Economist Intelligence Unit. (2008). *The Future of Enterprise Information Governance*. Londres, UK.
- Evora, V. (2009). *Procedimiento para la Gestión de la Calidad de Datos*. Las Villas: Universidad Central Marta Abreu .
- Fragoso, R. B. (2013). La Importancia de Gobernar Información en la Era de Big Data. *SG Software Guru*, 20.

- Geiger, J. (2004). *Data Quality Management The Most Critical Initiative You Can Implement*. Montreal.
- Gillies, A. (2011). Improving the Quality of Information Security Management Systems with ISO27000. *The TQM Journal*, 367-376.
- Giner, F. (2004). *Los sistemas de Información en la Sociedad del Conocimiento*. Madrid: ESIC.
- Gwen, T. (2015). *The GDI Data Governance Framework*. Londres.
- Informatica*. (2016, Junio 5). Retrieved from <https://www.informatica.com/services-and-training/glossary-of-terms/data-management-definition.html#fbid=6UsdqQn9VNx>
- Information Builders*. (2016, Junio 5). Retrieved from <http://www.informationbuilders.com/data-quality-management-dqm>
- Instituto Superior Nuestra Sra. de la Paz. (1016, Mayo 30). *Introducción a la Informática*. Retrieved from <https://datosuno.wordpress.com/>
- Juran, J. M. (1998). *Juran's Quality Handbook*. McGraw-Hill.
- Khatri, V., & Brown, C. (2010). Designing Data Governance. *Communications ACM*, 148-152.
- Lucero Ascencio, I. N. (2016). *Desarrollo de una Herramienta Data*. Barcelona.
- Moghe, P. (2009). Controlling Risk with a Data Governance Framework. *Bank Accounting & Finance*, 49-51.
- Moseley, M. (2008). Agile Data Governance: Key to Solving Enterprise Data Quality Problems. *Data Center Journal*, 24.
- Nazareth, D., & Choi, J. (2015). A System Dynamics Model for Information Security Management. *Information & Management*, 123-134.
- Needham, D., & Sinopoli, D. (2009). Improving Data Quality Control in Quality Improvement Projects. *Computers & Security*, 196-207.
- Negocio, S. e. (2016, Junio 15). *SINNEXUS Business Intelligence*. Retrieved from http://www.sinnexus.com/business_intelligence/piramide_negocio.aspx
- Otto, B. (2011). A morphology of the Organisation of Data. *ECIS*, 272.

- Panian, Z. (2010). Some Practical Experiences in Data Governance. World Academy of Science. *World Acad. Sci. Eng. Technol*, 150-157.
- Patrick, E. (2011). Effecting Data Quality Through Data Governance: a Case Study in the Financial Services Industry. *All Regis University Theses*, 465.
- Pennanen, I. (2014). *Gestión de Negocios Internacionales*. Turku, Finlandia: University of Applied Sciences.
- Power Data. Especialista en Gestión de Datos*. (2016, junio 10). Retrieved from <http://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/407376/Las-mejores-pr-cticas-de-data-governance-o-gobierno-de-datos>
- Prada, E. (2008). Los Insumos Invencibles de Decisión: Datos, Información y Conocimiento. *Anales de Documentacion*, 183-196.
- Revees, L. (2009). *A Manager's Guide to Data Warehousing*. Indianapolis, Indiana: Publicaciones Wiley.
- Ritchey, T. (2006). Problem Structuring Using Computer-Aided Morphological Analysis. *Journal of the Operational Research Society*, 792-801.
- SAS Latín América*. (2016, junio 10). Retrieved from <http://blogs.sas.com/content/sasla/2014/04/16/gobierno-de-datos/>
- Sheikhpour, R., & Modiri, N. (2012). An Approach to Map COBIT Processes to ISO/IEC 27001. *International Journal of Security and Its Applications*, 13-28.
- Stair, R. M., & Reynolds, G. W. (2010). *Principios de Sistemas de Información*. Santa Fe: Cengage Learning.
- Strout, S., & Eisenhauer, J. (2011). *The Elephant in the Room: Data*. published by The Data Governance Society.
- Susanto, H., & Nabil, M. (2012). Information Security Challenge and Breaches: Novelty Approach on Measuring ISO 27001 Readiness Level. *International Journal of Engineering and Technology*, 67-75.
- UCSC*. (2016, Agosto 16). Retrieved from <http://www.ucsc.cl/nuestra-universidad/direccion-de-admision-y-registro-academico/>
- UK, DAMA. (2013). *The Six Primary Dimension for Data Quality Assessment*. Reino Unido.

Wende, K. (2007). A Model for Data Governance – Organising Accountabilities for Data.
18th Australasian Conference on Information Systems, 417-425. Toowoomba.

Anexos

Anexo A. Catálogos de sistemas

A continuación, se detallan los sistemas que son maneja la DSI los cuales se dividen en sistemas del área académico curricular y del área administrativa.

- Sistemas del Área Académico Curricular: En la Tabla A-1 se detallan los sistemas y el grupo de usuarios que lo utilizan.

Tabla A-1 Sistemas del área académico curricular (Dirección de Servicios Informáticos, 2016).

SISTEMAS ACADÉMICOS					
Sistemas / Aplicaciones	Alumno	Ex Alumno	Docentes	Académicos	Administrativos
Aplicación Análisis Cuantitativo de Docencia Académica (ACDA)				X	X
Aplicación Carga Masiva de Datos					X
Aplicación CDinámico				X	X
Aplicación Certificados en Línea	X	X			X
Aplicación Ficha de Postulación Web	X	X		X	X
Aplicación Horario	X		X	X	X

Sistemas / Aplicaciones	Alumno	Ex Alumno	Docentes	Académicos	Administrativos
Aplicación Mensajería Masivo (Automail)				X	X
Aplicación Mis Cuentas	X	X		X	X
Aplicación Mis Registros / Información Estudiantes	X	X	X	X	X
Aplicación UCSC Móvil	X				
Plataforma de Enseñanza Virtual de Aprendizaje (EV@) Plataforma	X		X	X	
Inteligencia de Negocio DARA/Docencia/DGE (Cubix)			X	X	X
Sistema Académico Curricular (SIMBAD)				X	X
Sistema Carga Archivos Mineduc / CAE (Easys)					X
Sistema Compromiso de Desempeño Académico (SCDA)				X	X
Sistema Consolidación Docentes (SISCON)				X	X

Sistemas / Aplicaciones	Alumno	Ex Alumno	Docentes	Académicos	Administrativos
Sistema de CIDD (Centro de Innovación y Desarrollo Docente)			X	X	X
Sistema de Registro de Indicadores (SIGES)				X	X
Sistema Gestor de Encuestas				X	X
Sistema Gestión para Capacitación (OTEC)					X
Sistema Gestión Unidad DAE (Webdae)	X				X
Sistema Notas en Línea (Notas Parciales)	X		X	X	X
Sistema Solicitud de Contrato				X	X

- Sistemas del Área Administrativo: En la Tabla A-2 se detallan los sistemas y el grupo de usuarios que lo utilizan.

Tabla A-2: Sistemas del área administrativo (Dirección de Servicios Informáticos, 2016).

SISTEMAS ADMINISTRATIVOS					
Sistemas / Aplicaciones	Alumno	Ex Alumno	Docentes	Académicos	Administrativos
Sistema de CIDD (Centro de Innovación y Desarrollo Docente)				X	X
Autopayroll (Web) Portal de Proveedores					X
Portal de Reprogramación Individual FSCU		X			X
Sistema de Control de Recuperación de Licencias Médicas					X
Sistema de Formulación Presupuestaria				X	X
Sistema de Gestión Presupuestaria				X	X
Sistema Remuneraciones y Administrador de Personal (Payroll)					X
Sistema de Solicitudes Administrativas				X	X
Sistema Financiero Contable (SAU)					X
Sistema Fondo de Crédito Delfos		X			X

Sistemas / Aplicaciones	Alumno	Ex Alumno	Docentes	Académicos	Administrativos
Sistema Registro de Actividades Institucionales					X
Sistema de Registro de Convenios				X	X
Webdelfos 2.0 Fondo de Crédito (DELFOS)					X

Anexo B. Encuesta

Usando Google Forms como herramienta se generó una encuesta, la cual fue aplicada a 6 usuarios con distintos roles del sistema SIMBAD (secretaria, jefe de carrera o académicos). A continuación, se muestra el formato y las preguntas realizadas en la encuesta.

En cuanto a calidad y seguridad de los datos responda lo siguiente:

	Si	No aplica	No
1. ¿Al consultar a un alumno o a un docente en el sistema, se presentan campos vacíos?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. ¿Al realizar ingresos en el sistema, ya sean de solicitudes, datos de académicos o alumnos, existen campos de llenado no obligatorios?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. ¿Se presentan registros con nombre o direcciones cortadas en el sistema? (Nombre: María Ignacia Jav)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. ¿Al consultar por los datos de un docente, estos se encuentran siempre actualizados?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. ¿Los datos de los alumnos son consistentes? Ejemplo(País de residencia: China)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. ¿Los datos registrados de los docentes en el sistema, son concordantes con los datos que posee la Dirección de Gestión de Personal?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. ¿Al consultar distintos registros en el sistema, estos son desplegados con el mismo formato?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. ¿Es posible registrar en el sistema a un alumno con año de nacimiento 1900?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. ¿Al ingresar datos al sistema, independientemente del formato en que los ingrese, el sistema los corrige estandarizándolos todos?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. ¿Al consultar por un alumno en el sistema, es desplegado más de un registro?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. ¿Es posible generar una reserva de sala en el sistema, en un aula ya reservada?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. ¿Al consultar en el sistema existen el mismo número de matrículas que de registros de alumnos?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. ¿Los alumnos al realizar actualización de los datos en DARA, estos son actualizados en un periodo máximo de 2 días hábiles en el sistema?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. ¿Las notas de examen de título son registradas en el sistema en no más de 2 días hábiles?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. ¿Puede ingresar datos en el sistema como la fecha o el Rut en distintos formatos (dd/mm/aaaa o dd-mm-aaaa)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. ¿Se le permite ingresar al sistema Rut con más dígitos de los existentes (111.111.111-1)?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. ¿En el registro de personas en el sistema, campos como la dirección y el nombre son validados, evitando que se cometan errores de tipeo?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

18. ¿Firmó un acuerdo de confidencialidad para el acceso de datos personales y confidenciales en el sistema SIMBAD?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. ¿Puede acceder a datos en el sistema que le son innecesarios para su trabajo?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. ¿Tiene conocimiento sobre políticas o normativas para la seguridad de los datos en el sistema impuestas por la universidad?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21. ¿Puede modificar libremente los datos personales de un estudiante o académico sin una solicitud previa del individuo en cuestión?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. ¿Puede eliminar registros del sistema?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. ¿Al solicitar datos al sistema, ya sea de estudiantes o de académicos, estos siempre se encuentran disponibles?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. ¿Al solicitar datos al sistema, ya sea de estudiantes o de académicos, estos siempre se encuentran actualizados?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. ¿Los datos solicitados al sistema, se presentan de manera clara?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anexo C. Gráficos de las encuestas realizadas

A continuación, se detallan los resultados obtenidos luego de la realización de la encuesta.

En el Gráfico C-0-1 se muestra las distintas funciones los usuarios encuestados realizan en el sistema SIMBAD.

Marque una o más de las funciones que realiza en el sistema SIMBAD

(6 respuestas)

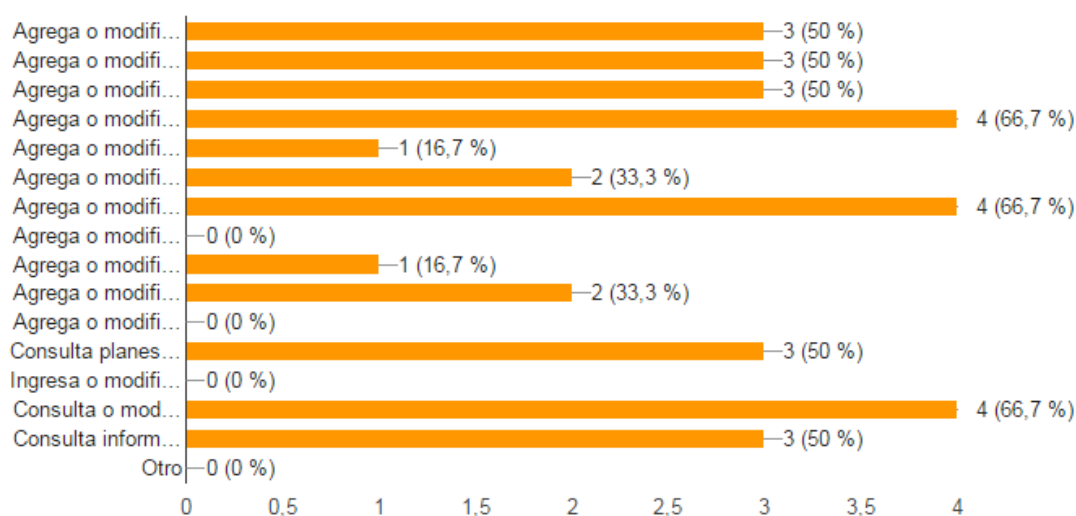


Gráfico C-0-1: Funciones realizadas por los usuarios en el sistema (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-2 las respuestas obtenidas en lo que respecta a completitud de los datos, detallando los resultados obtenidos.

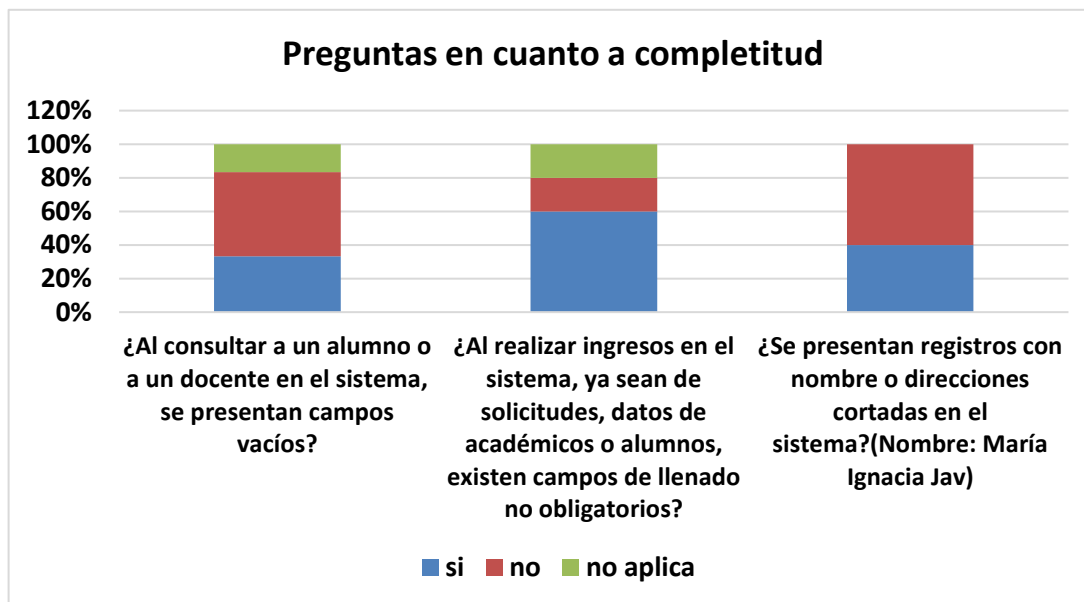


Gráfico C-0-2: Resultados por pregunta en completitud de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-3 las respuestas obtenidas en lo que respecta a la consistencia de los datos, detallando los resultados obtenidos.

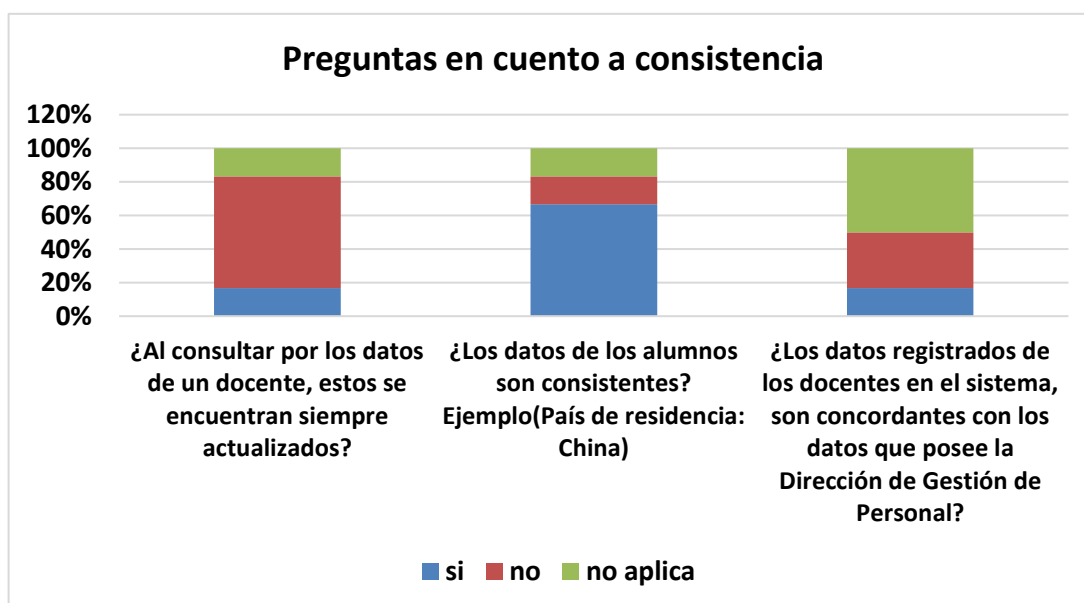


Gráfico C-0-3: Resultados por pregunta en consistencia de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-4 las respuestas obtenidas en lo que respecta a exactitud de los datos, detallando los resultados obtenidos.

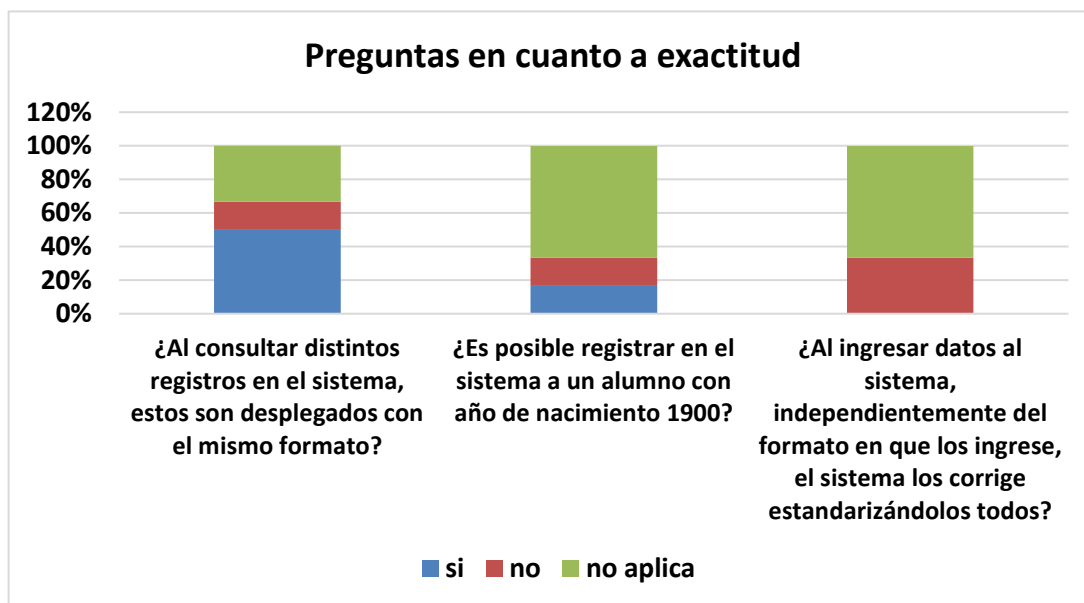


Gráfico C-0-4: Resultados en cuanto a exactitud de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-5 las respuestas obtenidas en lo que respecta a la unicidad de los datos, detallando los resultados obtenidos.

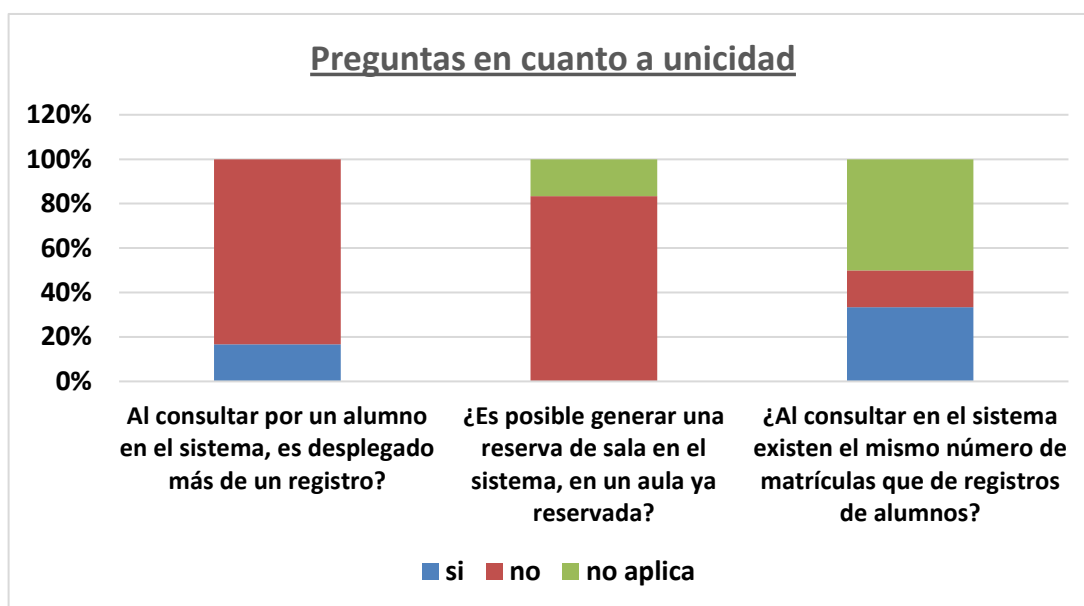


Gráfico C-0-5: Resultados por pregunta en unicidad de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-6 las respuestas obtenidas en lo que respecta a oportunidad de los datos, detallando los resultados obtenidos.

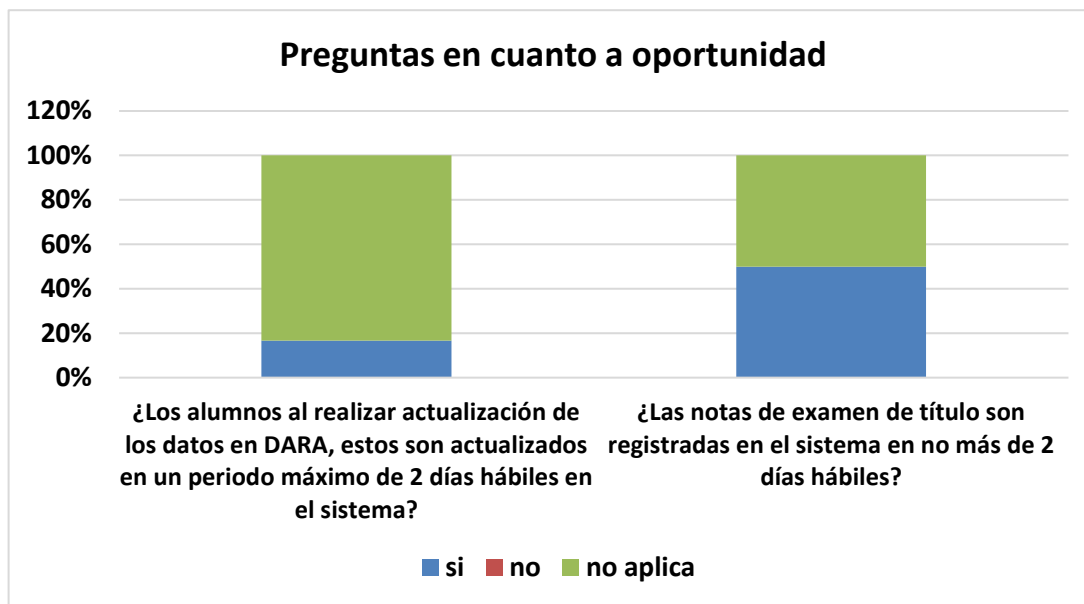


Gráfico C-0-6: Resultados por pregunta en oportunidad de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-7 las respuestas obtenidas en lo que respecta a validez de los datos, detallando los resultados obtenidos.

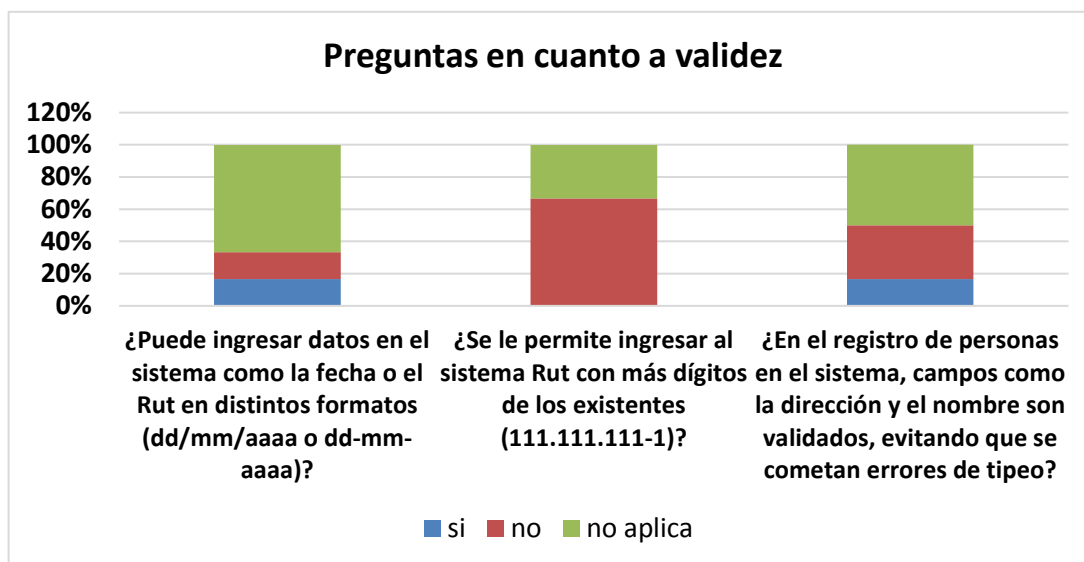


Gráfico C-0-7: Resultados por pregunta en validez de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-8 las respuestas obtenidas en lo que respecta a consistencia de los datos, detallando los resultados obtenidos.

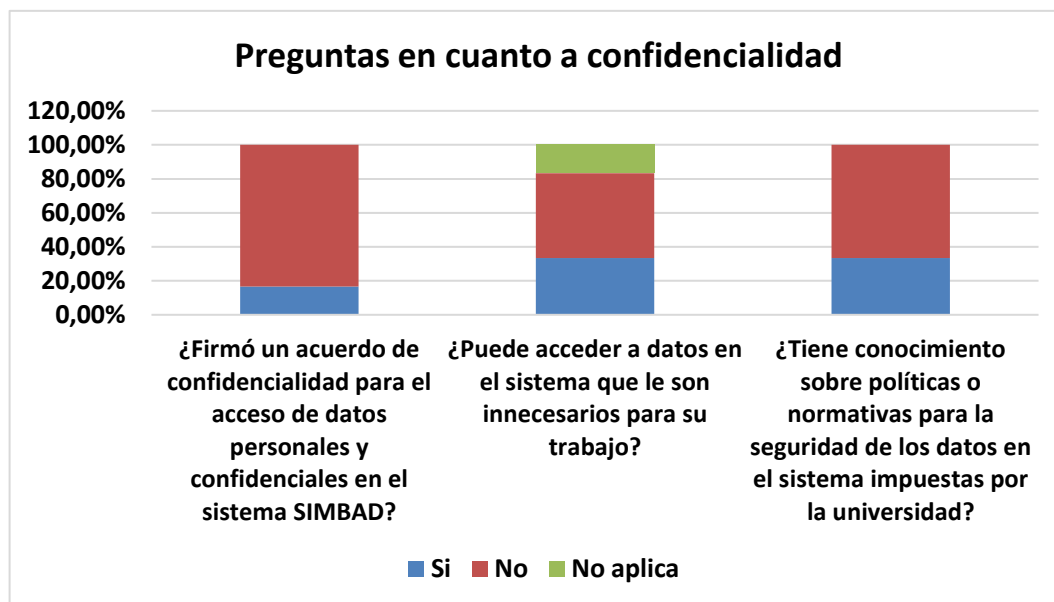


Gráfico C-0-8: Resultados por pregunta en confidencialidad de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-9 las respuestas obtenidas en lo que respecta a Integridad de los datos, detallando los resultados obtenidos.

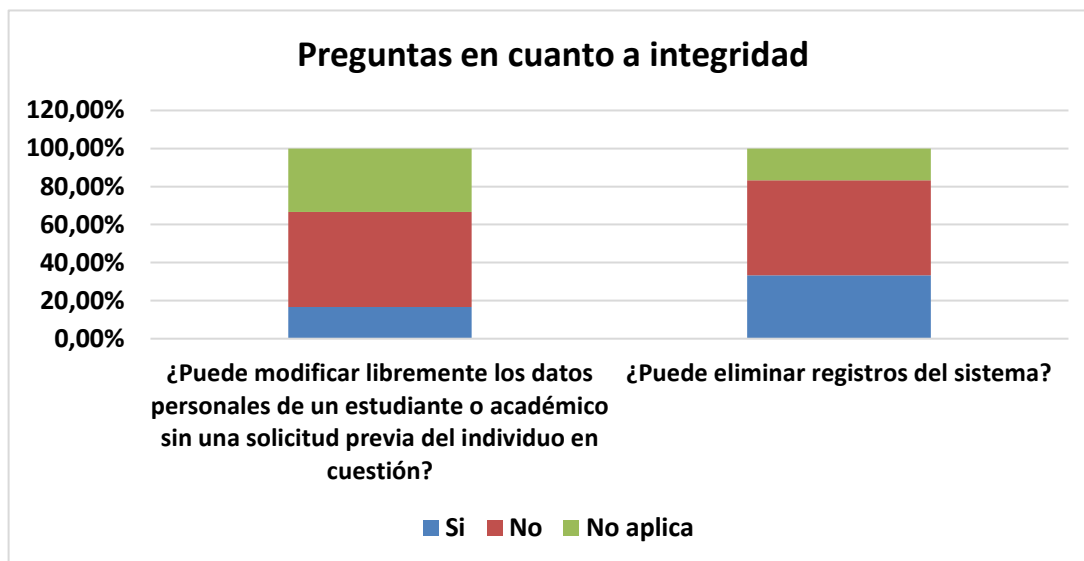


Gráfico C-0-9: Resultados por pregunta en integridad de los datos (Fuente: Propia).

Se pueden apreciar en el Gráfico C-0-10 las respuestas obtenidas en lo que respecta a la disponibilidad de los datos, detallando los resultados obtenidos.

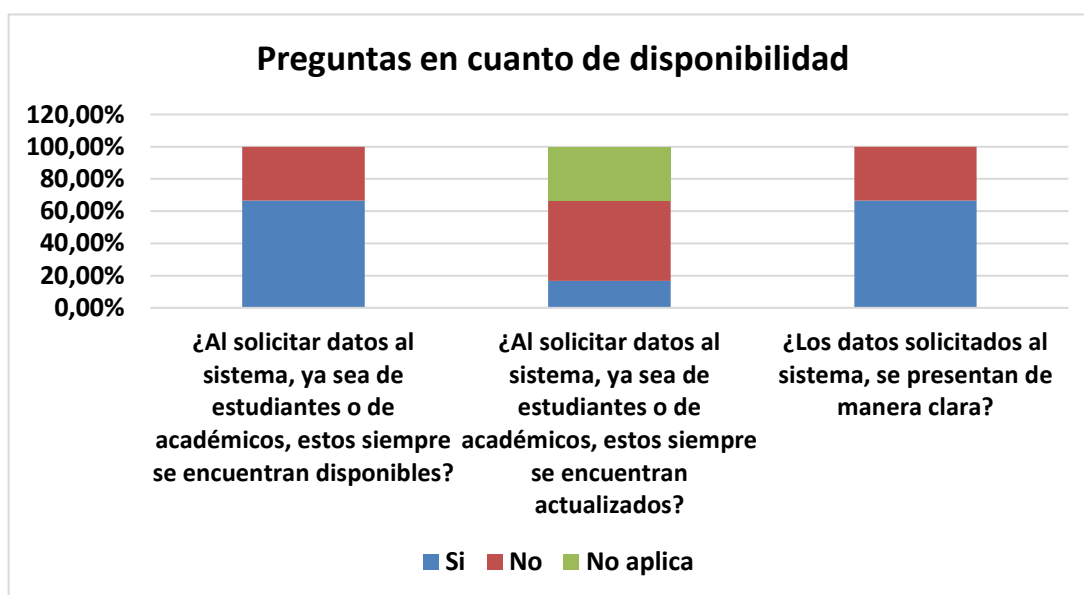


Gráfico C-0-10: Resultados por pregunta en disponibilidad de los datos (Fuente: Propia)

Anexo D. Datos activos para el sistema SIMBAD

A continuación, se detallan los datos activos de las tablas principales de la base de datos del sistema SIMBAD.

Tabla D-3: Tablas principales para el sistema (Fuente: Propia).

Tablas Principales	Datos activos
ARANCEL_HOMO	<p>A_H_C_SEM: Representa el código del semestre al cual se le realizará la convalidación.</p> <p>A_H_T_RAMO_HOMO: Representa el tipo de asignatura a la cual se realizará la convalidación.</p> <p>A_H_VALOR: Representa el monto del arancel por el que será la convalidación.</p>
CARRERAS	<p>CARR_CODIGO: Código específico que diferencia a una carrera de otra.</p> <p>CARR_DESCRIPCION: Nombre de la carrera.</p> <p>CARR_CAMPUS: Código del campus al cual pertenece la carrera.</p> <p>CARR_INST_UA_CODIGO: Unidad académica responsable de la carrera.</p>

Tablas principales	Datos activos
TIT_GRD:	<p>T_G_POND_PRES: Ponderación que se le otorga a la presentación de la defensa del título.</p> <p>T_G_POND_TRAB: Ponderación nota de tesis o proyecto.</p> <p>T_G_POND_EXAM: Ponderación nota de examen.</p>
PLANES_CARR	<p>PLAN_CODIGO: Código único que diferencia a una carrera de otra.</p> <p>PLAN_MAX_SEM: Semestre máximo cursado, se utiliza para validar las causales de eliminación.</p> <p>PLAN_MAX_INSC: Número máximo de cursos que el alumno puede inscribir.</p> <p>PLAN_CRED_MIN: Número mínimo de créditos que se debe aprobar en un semestre académico.</p>
PLANES_CARR_RAMO	<p>P_C_E_PLAN_CARR_RAMO: Indica la relación carrera ramo.</p>
CURSOS	<p>CURS_CODIGO: Código específico que diferencia a un curso de otro.</p> <p>CURS_CUPO: Cupo para el curso.</p> <p>CURS_C_SEDE: Código de la sede donde se imparte el curso.</p>

Tablas principales	Datos activos
RESTRIC_INSC	RI_CURS_CODIGO: Código del curso. RI_C_CARRERA: Carrera a la cual es ofertado el curso. RI_CUPO: Cupo del curso asignado a la carrera.
PROFESORES_CURSO	P_C_C_CURSO: Código del curso. P_C_PROF_RUT: Rut profesor asignado al curso.
HORAS_CURSOS	H_C_CURSO_CODIGO: Código del curso. H_C_HOR_CODIGO: Código y descripción del módulo asignado al curso H_C_N_SALAS_REQ: Salas requeridas para el curso. H_C_T_CLASE: Tipo de clase.
ACTA_CURSO	ACT_CURS_CODIGO: Código del curso. ACT_ESTADO: Estado del acta.

Tablas principales	Datos activos
CURSOS_INSCRITOS	<p>C_IN_CURSO_CODIGO: Código del curso.</p> <p>C_IN_SEME_ALUM_RUT_ALUM: Rut del alumno.</p> <p>C_IN_NOTA_FINAL: Nota final obtenida por el alumno en el curso.</p>
NOTAS_ACTI_ALUMNOS	<p>N_AA_ALUM_RUT: Rut alumno.</p> <p>N_AA_C_CURSO: Código del curso.</p> <p>N_AA_NOTA: Nota obtenida en la actividad.</p> <p>N_AA_N_ACTIV: Actividad realizada.</p> <p>N_AA_PROF_RUT: Rut profesor.</p>
HORAS_SALAS	<p>H_S_C_SALA: Código de la sala a reservar.</p> <p>H_S_F_INI: Fecha de Inicio de la reserva.</p> <p>H_S_F_FIN: Fecha de término de la reserva.</p> <p>H_S_CUR_COD: Código del curso que reserva la sala.</p> <p>H_S_PERS_RUT: Persona responsable por la reserva.</p> <p>H_S_H_C_HOR_COD: Módulo para el que se realiza la reserva.</p>

Tablas principales	Datos activos
PERSONAS	<p>PERS_RUT: Rut de la persona sin digito verificador.</p> <p>PERS_DV_RUT: Dígito verificador del Rut.</p> <p>PERS_NOM1: Primer nombre de la persona.</p> <p>PERS_NOM2: Segundo Nombre de la persona.</p> <p>PERS_AP_MATERNO: Apellido Materno de la persona.</p> <p>PERS_AP_PATERNO: Apellido paterno de la persona.</p> <p>PERS_E_MAIL: Correo electrónico de la persona.</p> <p>PERS_F_NACIMIENTO: Fecha de nacimiento de la persona.</p> <p>PERS_SUB_TIPO: Tipo de cargo de la persona en la universidad.</p> <p>PERS_C_EST_C: Estado actual de la persona en la universidad.</p> <p>PERS_SEXO: Sexo de la persona.</p>

Tablas principales	Datos activos
DIRECCIONES	<p>DIRE_CODIGO: Código específico que diferencia a una dirección de otra.</p> <p>DIRE_PERS_RUT: Rut de la persona a la que pertenece la dirección.</p> <p>DIRE_C_CIUDDAD: Ciudad a la que pertenece la dirección.</p> <p>DIRE_DIR1: Dirección primaria asignada a una persona.</p> <p>DIRE_FONOS: Teléfono asignado a una persona.</p>
RAMOS	<p>RAMO_C_RAMO: Código específico que diferencia a un ramo de otro.</p> <p>RAMO_NOMBRE: Nombre asignado al ramo.</p> <p>RAMO_C_DURA_RAMO: Duración que tiene el ramo.</p> <p>RAMO_C_RAMO_ESCUELA: Código nemotécnico que identifica el ramo.</p>
I_REQ	<p>IREQ_CODIGO: Código del requisito del ramo.</p> <p>IREQ_C_RAMO_REQ: Código del ramo requisito.</p>

Tablas principales	Datos activos
SALAS	<p>SALA_C_SALA: Código específico que diferencia una sala de otra.</p> <p>SALA_CAPACIDAD: Capacidad de la sala.</p> <p>SALA_UNIDAD: unidad responsable de la sala.</p> <p>SALA_UBICACION: Ubicación de la sala.</p> <p>SALA_DESCRIPCION: Descripción del tipo de sala.</p>