

ANTEPROYECTO: ESTUDIO SOBRE EL USO DE LA INFORMACIÓN NO
ESTRUCTURADA EN LAS PYMES DE DESARROLLO DE SOFTWARE DE LA
CIUDAD DE MEDELLÍN.

JULIAN GUZMAN ZAPATA

200029401010

jguzmanz@eafit.edu.co

JUAN FELIPE BORJA SANTAMARIA

200229419010

jborjasa@eafit.edu.co

ESCUELA DE INGENIERÍA UNIVERSIDAD EAFIT

UNIVERSIDAD EAFIT
ESCUELA DE INGENIERÍA
MEDELLÍN

2007

1 DEFINICIÓN DEL PROBLEMA

La información no estructurada es aquella que no proporciona un esquema con el cual se pueda realizar una interpretación para obtener información sobre su significado. La toma de decisiones en las organizaciones se fundamenta en información, la cual puede ser estructurada, semiestructurada o no estructurada. Usualmente el proceso de toma de decisiones se apoya en la información estructurada, y un alto porcentaje de información no estructurada no se tiene en cuenta para apoyar una decisión. Además, actualmente no existe una fuente de información o estudio que pueda dar testimonio del estado actual del uso y gestión de la información no estructurada en las Pymes de la ciudad de Medellín. Es así como se hace necesario determinar la situación actual acerca de la importancia y nivel de uso que le da la pequeña y mediana industria del entorno empresarial de la ciudad de Medellín a la información no estructurada.

2 OBJETIVO GENERAL

Identificar el estado actual del manejo y gestión de la Información no Estructurada, en las Pymes desarrolladoras de software ubicadas en la ciudad de Medellín.

3 OBJETIVOS ESPECÍFICOS

- Identificar la importancia de la información no estructurada en las organizaciones actuales.
- Diseñar una encuesta que permita describir el contexto del caso de estudio.
- Aplicar la encuesta y analizarla.

4 PRODUCTOS

- Documentación donde se incluye la información técnica de la encuesta y su respectiva explicación en el contexto, utilizando diagramas y gráficos estadísticos que faciliten su comprensión.
- Análisis de la encuesta.
- Conclusiones y recomendaciones.

5 BENEFICIARIOS

Todas aquellas personas que estén interesadas en la gestión de la información como punto de partida para la toma de decisiones y determinación de necesidades informáticas, tales como directivos de PYMES y analistas de información pertenecientes o no a estas. También se abre la posibilidad a Docentes y Estudiantes que deseen iniciar una nueva investigación tomando esta como punto de referencia.

Otros beneficiarios serían la industria del software, pues se detectarían necesidades y posibles nichos de mercado.

6 IMPORTANCIA DEL PROBLEMA DENTRO DE LA CARRERA

La información se ha convertido en un recurso vital para todo tipo de empresas. Tanto así, que a través de su uso se han logrado grandes mejoras competitivas, tanto en la toma de decisiones de una organización, como en el mejoramiento continuo de sus procesos. La información no estructurada, permite determinar necesidades internas y externas, pero requiere búsquedas precisas de información y el desarrollo de aplicaciones informáticas relacionadas. Es así pues, como un Ingeniero de Sistemas debe estar enterado de este tema en cuestión, ya que es un campo en su mayor parte inexplorado y de muchas aplicaciones específicas en cualquier sector empresarial.

7 METODOLOGÍA

En sus inicios, la investigación se basará en la consulta de la literatura para el estudio del tema de la información para luego realizar una profundización en la no estructurada. Esto incluye la lectura de textos, papers, proyectos de grado, artículos de revistas y posibles entrevistas. Posteriormente se utilizará este conocimiento como parámetro para diseñar una encuesta y realizar una investigación empírica y exploratoria de las pequeñas y medianas empresas. Luego de extraer estos datos de las organizaciones se hará un análisis y se utilizará una comparación para determinar la posición de estas empresas en materia de la utilización y gestión de la información no estructurada.

8 BIBLIOGRAFÍA

- DUEÑAS SANCHEZ, Henry. Análisis del impacto socio-cultural de las tecnologías de información y comunicación en Medellín, Colombia. Informe final de investigación. UNIVERSIDAD EAFIT. 2005.
- BENTLEY, Lonnie D.; WHITTEN, Jeffrey L.; BARLOW, Víctor M. Análisis y diseño de sistemas de información. MCGRAW-HILL / IRWIN 1996.
- GONZALEZ FERNANDEZ, Ciro. Características principales del diseño muestral general para encuestas multipropósito. U. DE LA HABANA: LA HABANA. 1987.
- HOYOS ALZATE, Carmen Lia; GARCIA GUZMAN, Eliana. Análisis crítico de herramientas para la evaluación de encuestas. MEDELLIN: UNIVERSIDAD EAFIT. 1992.
- AGUDELO B., Mónica Maria; GOMEZ O., Juan Carlos. Sistema para la toma de decisiones en investigaciones a partir de encuestas. MEDELLIN: UNIVERSIDAD EAFIT. 1997
- Biblioteca de administración de la pequeña y mediana empresa. MEXICO: MCGRAW-HILL. 1986.

9 PRESUPUESTO

- El presupuesto estará basado en los gastos que se generen a partir del transporte para realizar las encuestas y las fotocopias que se llevarán a cada empresa para ser respondidas.

Aproximadamente sería:

Transporte → Gasolina Corriente: 100.000 pesos

Fotocopias → 90 copias * 100 pesos para un total de 9.000 pesos

Lo que en total sería de 109.000 pesos m/l.

10 CRONOGRAMA

ACTIVIDAD	Febrero	Marzo	Abril	Mayo	Junio
	SEMANA				
Revisión Bibliográfica					
Marco Teórico					
Diseño de la encuesta, realización y recopilación.					
Análisis					

<p>Revisión de la propuesta con el asesor del proyecto para recibir sus sugerencias, y de acuerdo a ellas, hacer los cambios respectivos.</p>																								
<p>Conclusiones</p>																								
<p>Presentación final del proyecto.</p>																								
<p>Ajustes finales.</p>																								

Nombre del Asesor: Juan Guillermo Lalinde Pulido

Firma del Asesor: _____

**“ESTUDIO SOBRE EL USO DE LA INFORMACIÓN NO ESTRUCTURADA
EN LAS PYMES DE INGENIERIA DE SOFTWARE DE LA CIUDAD DE
MEDELLÍN”**

**JULIÁN GUZMÁN ZAPATA
JUAN FELIPE BORJA SANTAMARIA**

**Trabajo de grado para optar al título de
Ingeniero de Sistemas**

**ASESOR
Dr. JUAN GUILLERMO LALINDE PULIDO
DOCENTE DPTO. INFORMÁTICA Y SISTEMAS – EAFIT**

**UNIVERSIDAD EAFIT
DEPARTAMENTO DE INFORMATICA Y SISTEMAS
MEDELLIN
2007**

Nota de aceptación:

Firma del presidente del jurado

Firma del jurado

Firma del jurado

Medellín, 19 de Octubre de 2007

AGRADECIMIENTOS

Los autores desean agradecer a las siguientes personas que de una u otra manera colaboraron con la consecución de este estudio:

Al Señor Elkin Darío Medina Agudelo, Gerente General de MVM Ingeniería de Software S.A, muchas gracias. A todas aquellas personas que nos abrieron sus puertas y nos regalaron un poco de su tiempo para responder la encuesta. A nuestros padres, gracias por todo. A nuestras novias, por el tiempo que estuvieron solitas. A nuestro asesor, Juan Guillermo Lalinde P. y todas las otras personas que de una u otra forma nos colaboraron, ellos saben quienes.

TABLA DE CONTENIDO

1. INTRODUCCIÓN	1
2. PROBLEMA DE INVESTIGACIÓN	3
2.1 ANTECEDENTES	3
2.2 DEFINICIÓN DEL PROBLEMA	7
3 OBJETIVOS	9
3.1 OBJETIVO GENERAL	9
3.2 OBJETIVOS ESPECÍFICOS	9
3.3 METODOLOGÍA	10
4. MARCO TEÓRICO	11
4.1 DATOS	11
4.2 INFORMACIÓN	11
4.3 METADATOS	11
4.4 INFORMACIÓN ESTRUCTURADA	12

4.5 INFORMACIÓN SEMIESTRUCTURADA	13
4.6 LENGUAJES <i>MARKUP</i>	13
4.7 SGML	14
4.8 HTML	14
4.9 XML	14
4.10 INFORMACIÓN NO ESTRUCTURADA	15
5. INFORMACION NO ESTRUCTURADA EN LAS EMPRESAS	16
5.1 DOCUMENTOS	16
5.2 CORREOS ELECTRÓNICOS	17
5.3 AUDIO, VIDEO Y ANIMACIONES	19
5.4 IMÁGENES	20
5.5 CONOCIMIENTO	22
6. GESTIÓN DE LA INFORMACIÓN NO ESTRUCTURADA	23
7. PYME	23
8. SITUACION ACTUAL	24

9. DISEÑO DE LA ENCUESTA	27
9.1 JUSTIFICACIÓN	27
9.2 DISEÑO MUESTRAL	27
9.3 ELABORACIÓN DEL CUESTIONARIO	28
9.4 PERSONAS ENTREVISTADAS	32
10. PROCEDIMIENTO DE ADMINISTRACIÓN	33
11. ANÁLISIS	34
11.1 RESPUESTAS MÁS RELEVANTES SIN VARIABILIDAD	34
PREGUNTA 11.1.1	34
PREGUNTA 11.1.2	35
PREGUNTA 11.1.3	36
PREGUNTA 11.1.4	36
11.2 RESPUESTAS MÁS RELEVANTES CON VARIABILIDAD	37
PREGUNTA 11.2.1	37
PREGUNTA 11.2.2	38

PREGUNTA 11.2.3	38
PREGUNTA 11.2.4	39
PREGUNTA 11.2.5	42
PREGUNTA 11.2.6	43
PREGUNTA 11.2.7	44
PREGUNTA 11.2.8	46
PREGUNTA 11.2.9	46
PREGUNTA 11.2.10	47
PREGUNTA 11.2.11	48
PREGUNTA 11.2.12	48
PREGUNTA 11.2.13	50
PREGUNTA 11.2.14	51
PREGUNTA 11.2.15	52
PREGUNTA 11.2.16	53
12. RESUMEN	54

13. CONCLUSIONES	57
14. RECOMENDACIONES	62
BIBLIOGRAFÍA	64

LISTADO DE TABLAS

Tabla 1. Resumen de la encuesta.

54

LISTA DE FIGURAS

Figura 1. Periodicidad del mantenimiento del portal web	37
Figura 2. Resumen porcentual de Audio, video e imágenes.	40
Figura 3. Correlación de Video con audio e imágenes.	41
Figura 4. Correlación de Imágenes con audio y video.	42
Figura 5. Posibilidad de observar relaciones entre los contenidos.	48
Figura 6. Porcentaje estimado de Información Estructurada.	49
Figura 7. Porcentaje de toma de decisiones apoyadas en BD's.	50
Figura 8. Conocimiento acerca de la Información no Estructurada.	52
Figura 9. Porcentaje estimado de Información no Estructurada.	53

LISTA DE GRÁFICAS

Gráfica 1. Procesos de una Empresa Desarrolladora de Software. 28

1. INTRODUCCIÓN

No cabe duda que la información se ha reconocido como un recurso indispensable para el día a día de cualquier empresa. Una considerable cantidad de la información referente a los negocios de las empresas se almacena en bases de datos relacionales, facilitando la búsqueda y consulta de esta información. También es indiscutible que el crecimiento de estas bases de datos no se ha interrumpido desde que el costo de espacio para almacenamiento ha venido disminuyendo.

El proceso de toma de decisiones en una empresa se ve influenciado por información. En ella se basan las personas para seleccionar entre varias alternativas, la más conveniente. Las bases de datos relacionales tienen mucha importancia en este punto, ya que de estas se puede proveer de información de primera categoría a las personas. La información, en el contexto del presente estudio, se la clasifica de acuerdo con el conocimiento que se tenga de su semántica en tres categorías diferentes: estructurada, semiestructurada y no estructurada. Por ahora, resulta adecuado afirmar que la información que reside en las bases de datos relacionales es estructurada y que el resto pertenece a las otras dos categorías. Es decir, la información no estructurada la podemos encontrar en documentos de texto como correos u hojas de cálculo y también en objetos como imágenes, video y audio. Esta información es proporcionalmente, más abundante que la estructurada y su tasa de crecimiento es aún mayor. [DO029] y [KN038] Como dicen Imran R. Mansura y Sunita Sarawagi, “Los sistemas de Bases de datos son islas de estructuras en un mar de fuentes de datos no estructurados.” [SU028] Este tipo de información tiene la misma importancia que la estructurada en el proceso de toma de decisiones debido a que suministra elementos apropiados y significativos del negocio. Esto quiere decir, que si una empresa descarta la información no estructurada para la toma de decisiones, puede perder de vista

cantidades de contenidos con información valiosa acerca de una situación determinada, lo que representa una desventaja.

Una de las razones por las que aparentemente no se ven resultados en este campo es porque el manejo de la información no estructurada es un problema complejo, inclusive en el manejo de textos donde a primera vista el hecho de analizar documentos en búsqueda de su significado, puede parecer un problema trivial. Para resaltar más esta situación, algunas empresas para descubrir y extraer conocimiento tienen en producción algunos sistemas de gestión documental con el fin de mantener todos los contenidos controlados y organizados. Pero el conocimiento actual acerca de los sistemas de bases de datos, y en particular la facilidad que se tiene de realizar consultas específicas, puede hacer emitir la afirmación precipitada de que la información estructurada es la que estamos acostumbrados a administrar y a procesar para dar soporte a la toma de decisiones. En este trabajo se tuvo por objetivo observar si las empresas de desarrollo de software de la ciudad de Medellín están utilizando la información no estructurada y cuál es el tratamiento que estas empresas le están dando a la misma. Estos datos provienen de la mejor técnica para obtener datos cualitativos y cuantitativos: entrevistas acompañadas de una encuesta guía que se le realizó a una muestra significativa de las empresas de la ciudad de Medellín. Se pretende así, recoger datos acerca de los diferentes tipos de información no estructurada que se manejan a nivel organizacional (audio, video, imágenes y texto incluyendo el correo electrónico) y cuáles son las áreas, cargos involucrados y métodos utilizados para su manejo. Además, se pudo evaluar cuál es el nivel de conocimiento que las empresas tienen de la información no estructurada, el uso de herramientas para la gestión de ésta y datos porcentuales sobre el uso de los diferentes tipos de información en los procesos de tomas de decisión.

2. PROBLEMA DE INVESTIGACION

2.1 ANTECEDENTES

En la década de los setentas se dieron grandes avances para proveer a los datos de una vista lógica. Estos avances se vieron representados principalmente por la aparición de modelos de datos, a saber: El modelo jerárquico, el modelo de red, el modelo relacional y el modelo de entidad-relación. [ML026]

El primer modelo que surgió fue el modelo jerárquico., en dónde “el esquema es una estructura arborescente compuesta de nodos, que representan las entidades, enlazados por arcos, que representan las asociaciones o interrelaciones entre dichas entidades”. [CA001]

El modelo de red, para proporcionarle estructura a los datos, básicamente introduce la separación diferenciada entre las entidades y las relaciones, en donde se representan las entidades en forma de nodos de un grafo, y las relaciones entre éstas mediante arcos que unen dichos nodos. Este modelo surge para solucionar problemas específicos del modelo jerárquico, tales como: hijos con más de un padre, interrelaciones reflexivas, más de una interrelación entre dos entidades e hijos sin padre. [CA001]

Por otro lado, el modelo relacional está basado en la teoría relacional y considera la base de datos como una colección de relaciones, en donde cada una de ellas se representa mediante una tabla, cada fila representa una

colección de valores que describen una entidad del mundo real, denominada tupla o registro, y cada columna se denomina campo.

Los modelos jerárquico, red y relacional proponen como realizar una abstracción del modelo físico utilizado por la base de datos pero tienen como punto de partida la representación de la información y no el dominio del problema que se va a trabajar. El modelo entidad relación, que originalmente fue propuesto como una alternativa más para abstraer el modelo físico, resultó ser adecuado para modelar partiendo del dominio del problema y fue modificado y adaptado para este fin. [MK012]

Las bases de datos relacionales utilizan el modelo relacional, pero la modelación conceptual se realiza utilizando el modelo entidad-relación. Este modelo, consta de conceptos básicos como entidades, relaciones y atributos. Las entidades son aquellos objetos acerca de los cuales se quiere almacenar información en la base de datos. Una relación es la correspondencia o asociación entre entidades. Mientras que un atributo es una propiedad o característica que tiene una entidad o relación. [CA001]

El modelo entidad relación¹ es el de más importancia en el tema que concierne a este estudio, el cual reúne el mayor número de ventajas con respecto a los de los modelos anteriormente mencionados, debido a que adopta el punto de vista más natural para hacer el modelo consistente con la realidad, a saber: entidades y relaciones. Además, se basa simultáneamente en la teoría de conjuntos y la teoría de relaciones.

¹ Las bases de datos relacionales trabajan de acuerdo con el modelo relacional, mientras que el modelado conceptual se hace mediante el modelo entidad-relación. Nos centramos en el modelo entidad-relación porque el problema trabajado pertenece al plano conceptual.

Las empresas han tomado este modelo de datos para representar su información. Como dice C. Wood acerca de la importancia de la información en las empresas: “La información es un recurso crítico en las empresas actuales, ya sean de tipo industrial, comercial o civiles. Las empresas están automatizando no solo las funciones básicas operacionales, sino también funciones de soporte administrativas...” [WO009]

En el contexto actual se define información estructurada como aquella que tiene un esquema que describe los datos y que proviene de un modelo de datos. El ejemplo más común de este tipo de información es la que se almacena en las Bases de Datos. Esta información está estrechamente relacionada con la toma de decisiones, como lo afirma Carlos Obeso, “la toma de decisiones en las organizaciones se fundamenta en información. Solo una parte pequeña de esa información esta estructurada, no más del diez por ciento. La información estructurada es muy útil pero es limitada para las necesidades de toma de decisiones complejas”. [OC011]

Así mismo J. March se refiere a este mismo tema de la siguiente manera: “Cuando se discute el diseño de los sistemas de información en las organizaciones, el valor de la información se enlaza ordinariamente a la gestión de una manera simple. La gestión se encarga primordialmente de tomar decisiones y se escoge entre varias alternativas basándose en la información acerca de la consecuencias y preferencias que son condicionales en una decisión”. [MA010]

La información no estructurada se define como aquella que no tiene un esquema que describe los datos y por lo tanto no se puede obtener una interpretación acerca de su significado. Carlos Obeso aclara que “en las organizaciones existe una enorme cantidad de información no estructurada”

[OC011] en forma de videos, audio, documentos escritos, imágenes, e incluso conocimiento tácito que reside en el intelecto de las personas. Esta es una de las razones que motivan este estudio, ya que varias investigaciones estiman que el 80% de toda la información corporativa es no estructurada. [FE006]

Como se mencionó anteriormente, la toma de decisiones se basa en información estructurada, pero este mismo autor afirma que también se basa implícitamente en información no estructurada, aunque no se extrae toda su riqueza oculta por su naturaleza.

2.2 DEFINICIÓN DEL PROBLEMA

Debido a que los estudios revelan que el 80% de la información que se maneja en las empresas es no estructurada, el trabajo se enfoca en analizar los diferentes factores involucrados por esta afirmación. Es por esto, que se estableció cuáles son los diferentes tipos de información no estructurada que se utilizan en las empresas de desarrollo de software de la ciudad de Medellín. Así mismo, se hace relevante conocer cuál es conocimiento que se tiene acerca de la información no estructurada, cómo se utiliza y cómo se apoyan en ésta los empresarios para la toma de decisiones. Cabe resaltar, que para ello, se evaluó el uso de información estructurada, a través de las bases de datos, para identificar qué porcentaje de ésta se maneja y de igual forma, en qué porcentaje se toman de decisiones basándose en este tipo de información.

La información en los últimos años ha tomado tanto valor que inclusive se ha dicho que la información es poder. [VI014], [MA040] y [WE039] Tanto es así: "Que cada vez hay un mayor número de organizaciones que predisponen todos los medios necesarios para su obtención y control". [GI011] Además, continúa el autor, muchos directivos dedican una alta parte de su tiempo a acceder a más información. Todo esto con un solo fin: "Que les conduzca a mejorar su proceso de toma de decisiones".

Charles H. Kepner aclara este punto de la siguiente manera:

“La información es la materia prima de la dirección. Es lo único con lo que cuenta un directivo para trabajar: Información acerca del mundo que lo rodea, de su organización y de sus planes, del funcionamiento de aquel segmento de la organización en el cual él es el responsable de

llevar a buen término estos planes, de las personas, las cosas y las condiciones. Tiene que saber qué información posee sobre cualquier problema, cuál no posee y cómo puede obtenerla, y cómo puede utilizar toda la información de que dispone, de la manera más ventajosa, a fin de hacer que el problema se resuelva. Debe ser tan perspicaz para reconocer, antes de lanzarse a actuar, qué información es aplicable e importante y cuál no lo es, como lo es para reconocer, después de ocurridos los hechos, la importancia evidente de cierta información”.
[KE012]

Además, aunque se ha realizado una búsqueda exhaustiva, no se ha podido encontrar una fuente de información o estudio que pueda dar testimonio del estado actual del uso y gestión de la información no estructurada en las Pymes de la ciudad de Medellín en el sector de desarrollo de software. Es así como se hace necesario determinar la importancia y nivel de uso que le da la pequeña y mediana industria del entorno empresarial de la ciudad de Medellín a la información no estructurada

3 OBJETIVOS

3.1 OBJETIVO GENERAL

Analizar cómo gestionan y cómo usan la información no estructurada en la toma de decisiones las Pymes de la ciudad de Medellín en el sector de desarrollo de software.

3.2 OBJETIVOS ESPECÍFICOS

1. Determinar el tamaño de las muestras necesarias para que el estudio sea significativo.
2. Establecer la proporción de uso de la información no estructurada en las empresas del sector objeto de estudio.
3. Determinar las formas de uso más comunes de la información no estructurada en las empresas del sector objeto de estudio.
4. Identificar las principales herramientas que se utilizan actualmente para el uso de la información no estructurada en las empresas del sector objeto de estudio.
5. Determinar el tipo de información no estructurada existente en las Pymes del sector objeto de estudio en Medellín
6. Determinar si las Pymes del sector objeto de estudio en Medellín gestionan la información no estructurada.
7. Determinar la influencia de la Información no estructurada en la toma de decisiones de las Pymes del sector objeto de estudio en Medellín.

3.3 METODOLOGÍA

La investigación se basó en la consulta de la literatura para el estudio del tema de la información para luego profundizar en la información no estructurada. Esto incluyó la lectura de textos, papers, proyectos de grado, artículos de revistas y entrevistas. Posteriormente se utilizó este conocimiento para diseñar una encuesta y realizar una investigación empírica y exploratoria de las pequeñas y medianas empresas. Luego de extraer estos datos de las organizaciones se hizo un análisis y se comparó la posición de estas empresas frente a la utilización y gestión de la información no estructurada.

4. MARCO TEÓRICO

Con el fin de establecer un marco de referencia común, se presentan a continuación las definiciones básicas que se utilizan en la construcción del marco teórico.

4.1 DATOS

“Dato es una colección de hechos considerados de forma aislada. Los datos describen una organización. Estos hechos aislados portan un significado, pero en general no son de utilidad por sí solos.” [WH002]

4.2 INFORMACIÓN

“Información es un dato que ha sido manipulado, con lo que resulta de utilidad para alguien. En otras palabras, la información debe tener valor, o en caso contrario sería un dato. La información dice a la gente algo que no sabía o le confirma algo que sospechaba.” [WH002] Esta afirmación quiere decir que la información tiene valor y es útil para alguien ya que está compuesta por hechos que están dentro de un contexto determinado.

4.3 METADATOS

La definición más simple de Metadatos es datos acerca de los datos. “Es información acerca de la organización de los datos, su dominio y las relaciones entre ellos”. [BA005] Un Metadato puede describir un dato o una colección de los mismos. Se utiliza para facilitar el entendimiento, uso y gestión de los datos. Los Metadatos varían dependiendo del tipo de datos a describir y del contexto en el cual se usan.

4.4 INFORMACIÓN ESTRUCTURADA

Es la información que ha sido analizada y organizada. Tiene un esquema que describe los datos y que proviene de un modelo de datos. Por lo general se almacena en una base de datos, motor de búsqueda o base de conocimiento. Como afirma David Ferrucci y Adam Lally “La información estructurada está caracterizada como la información cuyo significado intencional no es ambiguo y está explícitamente representado en la estructura o formato de los datos. El ejemplo canónico de la información estructurada es la tabla de la base de datos relacional” [FE006] Cada elemento de la información está asociado con un esquema preciso y bien definido donde cada encabezado de una columna de la base de datos especifica su semántica, definiendo cómo la información debería ser interpretada por una aplicación informática o por un usuario final.

Se mencionó anteriormente que la información está compuesta por hechos y, al relacionarlo con las representaciones, M. Senko afirma que “La información consiste de hechos acerca de las cosas. Estos hechos y cosas existen independientemente de cualquier representación, pero es esencialmente imposible operar conceptualmente con ellos excepto en términos de alguna representación”. [SE008] Acerca de estas representaciones, M. E. Senko especifica que “las representaciones son materializaciones concretas de la información estructurada. [...] Los computadores trabajan directamente con las representaciones y sólo indirectamente, si mucho, con la información representadas por ellas.” [SE008]

4.5 INFORMACIÓN SEMIESTRUCTURADA

La información semiestructurada es aquella en donde los datos no se rigen por un esquema estricto de una base de datos, es decir, su estructura no es regular. Son datos que pueden ser irregulares o estar incompletos y tienen una estructura que puede ser variable. La información semiestructurada está constituida en esencia, por documentos, tales como documentos de texto y presentaciones en diferentes formatos. “Tal información es típicamente semiestructurada, esto es, que existe alguna estructura en los documentos pero no exactamente una estructura formal impuesta por un esquema de bases de datos”. [MA004]

Según Ke Wang y Huiqing Liu: “La información semiestructurada se presenta cuando la fuente de los datos no impone una estructura rígida (tal como la Web) y cuando los datos provienen de fuentes heterogéneas. A diferencia de la información no estructurada, la semiestructurada si tiene algún tipo de estructura.” [WA003]

4.6 LENGUAJES *MARKUP*

Los lenguajes *Markup* [BA005] son aquellos que especifican información acerca del texto, como el formato y la presentación, la estructura o la semántica del mismo. Esta información se define por medio de marcas (de ahí su nombre *Markup*) denominadas etiquetas, que por lo general se presentan, una al inicio de un metadato y otra al final de este. Las etiquetas además, cuando están bien manejadas, garantizan la elaboración de un documento bien definido y formado. Entre los lenguajes de *markup* más importantes se encuentran SGML, HTML y XML.

4.7 SGML

El Estándar Generalizado de lenguaje de marcas, SGML por sus siglas en inglés *Standard Generalized Markup Language*, provee las reglas para definir un lenguaje de marcas. El SGML está basado en un trabajo previo realizado por IBM denominado GML, cuyas siglas significan *Generalized Markup Language*. [BA005]

4.8 HTML

El lenguaje de marcas de hipertexto (*HyperText Markup Language*), es una instancia de SGML. Fue creado para almacenar los documentos Web. Hay que subrayar que el HTML no permite a los usuarios crear etiquetas o atributos específicos para darle una semántica determinada a sus datos porque contiene etiquetas predeterminadas. [BA005]

4.9 XML

XML significa lenguaje de marcas extendido (eXtensible Markup Language). A diferencia del HTML, XML no es una instancia de SGML, sino que es un subconjunto simplificado del mismo. Con este lenguaje es posible definir formatos específicos para documentos a partir de etiquetas propias del usuario. Además, impone una sintaxis rígida la cual es importante para el procesamiento y transmisión de datos. [BA005]

4.10 INFORMACIÓN NO ESTRUCTURADA

Son diferentes las definiciones en este campo, pero cuando se refiere a información no estructurada, esencialmente se trata de todos los contenidos almacenados de los cuales no se puede decir específicamente cómo están organizados, cómo están relacionados con el resto de la información o cómo deben ser usados. En esencia, hacen parte de esta categoría los objetos tipo mapa de bits, los cuales no están basados en ningún lenguaje natural (imágenes, video y audio) y los archivos de texto, que están basados en lenguaje escrito, tales como el correo electrónico. En esta categoría se clasifican todos los documentos para los cuales no hay metadatos semánticos.

La información estructurada puede ser manejada con soluciones que soportan “*queries*” y reportes a tipos predeterminados de datos y relaciones previstas. La información no estructurada no tiene una definición conceptual ó de tipo de datos y para efectuar consultas hay que buscar métodos automáticos para asociarle alguna estructura. Actualmente lo más utilizado es una aproximación sintáctica, i.e. búsqueda de palabras, o utilizar técnicas *inteligentes* tales como el procesamiento de lenguaje natural, redes neuronales, etc.

5. INFORMACION NO ESTRUCTURADA EN LAS EMPRESAS

5.1 Documentos

En las empresas se pueden observar dos clases de documentos: Electrónicos que son todos los que están en un archivo digital (Word, Excel, pdf, etc.) y tangibles que son todos aquellos documentos que están en papel, ya bien porque han sido imprimidos o tiene su origen en un formato físico (notas, balances generales, hojas de vida, etc.). Los archivos en papel todavía tienen una gran implementación en las organizaciones. Mucha parte de la información sigue siendo archivada de manera tangible generando problemas a la hora de su gestión no sólo por la dificultad de realizar búsquedas sobre ellos, sino también por el espacio que estos ocupan.

A pesar de que cada día se están generando más documentos electrónicos, “la acumulación de documentos electrónicos en los ordenadores de las empresas aparece como un nuevo problema que hasta ahora no existía. El espacio en este caso no preocupa demasiado, pues la memoria de almacenamiento es cada vez más barata; pero la dificultad para encontrar lo que se busca y la generación de copias de seguridad empiezan a ser considerados problemas importantes de gestión.”[BC015]

Claro que esto no sucede solamente para los documentos electrónicos. En el caso de los archivos físicos también se generan grandes problemas a la hora de querer conseguir un documento específico porque pueden surgir inconvenientes al buscar y relacionar los diferentes documentos, especialmente cuando el archivo crece demasiado.

Sin embargo, se hace indispensable que para ambos casos exista una gestión documental que sirva de apoyo a las personas y al negocio como tal. Los objetivos básicos se pueden definir de la siguiente manera:

- “Facilitar a las personas trabajar con los documentos. Cada persona sabe qué documentos tiene que guardar, cuando, como y donde. Cada persona sabe como encontrar en poco tiempo los documentos adecuados cuando los necesita.
- Facilitar que la información se comparta y se aproveche como un recurso colectivo, evitar que se duplique, evitar fotocopios innecesarios, evitar dobles grabaciones de datos, etc.
- Conservar la memoria de la organización más allá de los individuos que trabajan en ella y poder aprovechar el valor de los contenidos en los que queda plasmada la experiencia, evitando empezar de cero sobre aspectos en los que ya hay experiencia acumulada.” [BC015]

5.2 Correos Electrónicos

Microsoft, en su página web española, destaca: “Sin lugar a dudas, Internet se ha convertido en la herramienta de comunicaciones empresariales más utilizada. Según algunos cálculos, su inmensa popularidad ha generado 275 millones de buzones corporativos. Otras estadísticas indican que el 60 por ciento de la información empresarial más importante se transmite en la actualidad por correo electrónico”. [MI016]

Es tal el impacto que el correo electrónico ha generado en las empresas, que puede volverse bastante peligroso debido a que mucha información importante está viajando por este medio y puede ser robada en cualquier momento.

Además, se ha llegado a ver, que los mismos correos electrónicos pueden ser utilizados como pruebas en las cortes judiciales. [AP032]

Otro elemento adicional es que, si no existe una disciplina de almacenamiento y gestión de los correos, no queda registro histórico de la información que fluye. Si las comunicaciones formales de la organización están basadas en el correo electrónico, se pierde la historia de las decisiones. “Los comunicados electrónicos y el intercambio de la información constituyen el vector de pérdida de información más importante para las empresas.” [GI033]

Según una encuesta realizada por Harris Interactive (2003): “El email es la tecnología favorita seguida del teléfono móvil, los portátiles, el teléfono, el procesador de textos, la hoja de cálculo, el fax y otros.” [UN018] Es indispensable comentar, según la misma encuesta, que el 83,7% de las empresas tienen como herramienta el correo electrónico.

Esta tecnología ha sido muy acogida por los empresas ya que involucra un bajo costo, tiene una amplia difusión (clientes, proveedores, público, empresas, etc.), facilidad de uso, diversidad de usos (envío de archivos, agenda, búsqueda, etc.). [UN018] Además permite la comunicación rápida y a muy bajo costo con cualquier persona conectada a Internet, independiente de su ubicación física. Más aún, a pesar de carecer de estructura, su implantación es inmediata y se puede utilizar fácilmente.

Pero es ahí dónde vienen los problemas: Al generarse un volumen de información bastante alto por usuario, se hace difícil rastrear los correos y por tanto encontrar, en muchos casos, información relevante que se necesita de manera oportuna. Por otro lado, no todos los correos que se transmiten están

relacionados con las funciones que desempeña el usuario. También circulan correos que contienen chistes, pornografía, circulación de mensajes, es decir, “cadenas de correo electrónico, que consisten en reenviar un mensaje a mucha gente” y que causan gran pérdida en cuanto a la productividad se refiere, ya que “se considera que el 32 por ciento de los mensajes leídos y enviados carecen de relevancia.” [IT034]

Por todo esto, es importante llevar un control de los correos electrónicos implantando políticas para el uso del correo electrónico en el trabajo, así como una serie de principios que regulen el uso de esta tecnología.

5.3 Audio, Video y Animaciones

“El contenido de video está proliferando en las empresas e incluye video conferencias grabadas, transmisiones por la Web, capacitación de los empleados, direcciones de los ejecutivos, y más. Las empresas se están esforzando para lidiar con este volumen de contenido cada vez mayor mientras mejoran el video como un factor estratégico. Los productos de puntos han forzado tradicionalmente a los clientes a unir las soluciones separadas”. [MK019]

El video en las organizaciones puede ser utilizado de diferentes formas. Una manera de usarlas es, las videoconferencias, que se realizan con personas ubicadas en diferentes lugares geográficos. Así mismo, puede utilizarse para concretar negocios y validar propuestas. También se utilizan en: relaciones con inversionistas, anuncios internos, reuniones de empleados, etc. En la mayoría de casos, estas videoconferencias no suelen quedar archivadas: se realizan en vivo sin generar ningún almacenamiento para su posterior utilización. Por otro

lado, el video ayuda en las capacitaciones que se le dan a los empleados al ingresar a las empresas, e inclusive, para generar conocimiento de una herramienta dada. Esto se puede evidenciar mediante la extracción del conocimiento a los empleados, aprendizaje a distancia, conferencias grabadas, programas de estudio individual e inclusive certificaciones para los empleados. El contenido de video también puede ayudar a reducir los costos de soporte del cliente al proporcionar detalles de los productos, instrucciones de instalaciones, guías para ubicación y resolución de problemas y manuales de usuarios.

Cabe destacar que en este mismo campo, también están las animaciones, las cuales son “una simulación de movimiento producida mediante imágenes que se crearon una por una; al proyectarse sucesivamente estas imágenes (denominadas cuadros) se produce una ilusión de movimiento, pero el movimiento representado no existió en la realidad.” [AN035] Éstas cumplen un papel similar en las empresas. Por lo general, las animaciones se crean utilizando herramientas como FLASH² y sirven de soporte para las capacitaciones, presentaciones de productos, campañas de mercadeo, promociones, etc. ya sea en la página Web o para vender productos directamente.

5.4 Imágenes

Son diversos los formatos de imágenes que se manejan a nivel corporativo y dependen en gran medida de la unidad de negocios que cada empresa maneja dentro de su portafolio de productos o servicios. Es diferente la organización, mantenimiento y gestión que le pueda dar una organización de Diarios Nacionales a la que pueda darle una Empresa Promotora en Salud. Sin

² Adobe Systems Incorporated.

embargo, para ambos es de vital importancia la información que allí pueda generarse.

Además de todos los formatos que se manejan (.jpg, gif, .bmp, .png, .tiff, etc.)³ existen otras, no electrónicas, que también son importantes para las empresas. Una imagen puede ser: una radiografía, un caso de uso, un logo en papel, un mapa e inclusive una señal de tránsito. Todas ellas cumpliendo un papel fundamental para cada organización. Otros usos de las imágenes son: el logo de la empresa, imágenes para capacitaciones, explicaciones de productos, mercadotecnia, etc. [IA020]

Cabe resaltar que a medida que transcurre el tiempo, se hace de mayor importancia tener todas las imágenes en medio electrónico. Esto se debe en gran medida, a que las imágenes tangibles pueden ser alteradas afectando su estado de conservación, su manipulación es lenta (debido a que implementar sistemas de calidad e intentar recopilar los originales se convierte en una tarea más dificultosa) y presentan problemas para su traslado.

Para el caso de las empresas de desarrollo de software, gran parte del conocimiento que les permite ser competitivos en el negocio está sustentado en diagramas. Para esto se utilizan diversas herramientas que ayudan a su elaboración pero no a su organización. Tampoco podemos dejar de lado todas aquellas imágenes que aún se desarrollan y archivan en papel.

³ JPEG: (Joint Photographic Experts Group), Creado por Joint Photographic Experts Group.
PNG: (Portable Network Graphics), Formato aprobado por el W3C.
GIF: (Graphics Interchange Format), Creado por CompuServe.
BMP: (BitMaP) Formato estándar de Windows. Por Microsoft Inc.
TIFF: (Tagged Image File Format), Fue desarrollado por Aldus y Microsoft.
Image file formats en http://en.wikipedia.org/wiki/Image_file_formats

5.5 Conocimiento

En las empresas existe otra fuente de información no estructurada, la cual reside en el conocimiento de las personas. Al respecto, Juan Carrión afirma que:

"El Conocimiento es un conjunto integrado por información, reglas, interpretaciones y conexiones puestas dentro de un contexto y de una experiencia, que ha sucedido dentro de una organización, bien de una forma general o personal. El conocimiento sólo puede residir dentro de un conocedor, una persona determinada que lo interioriza racional o irracionalmente". [CO021]

Este mismo autor nos afirma que este conocimiento influye en el diario accionar de las personas, tomando decisiones acerca de los fenómenos percibidos, además, que es uno de los intangibles que permite obtener una ventaja competitiva. Para este último punto, Francisco Martínez aclara que "el conocimiento se transforma en un aprendizaje continuo que es vital para el funcionamiento sostenible de las organizaciones" [MA022] relacionado con las competencias, con la capacidad de innovar y mejorar, con el compromiso y la motivación.

6. GESTIÓN DE LA INFORMACIÓN NO ESTRUCTURADA

Según Sudhakar Muddu, “la información no estructurada es información que por lo general no ha sido gestionada. Este tipo de información se almacena de diversas formas y se crean individual y manualmente más que automáticamente. El sistema de archivos en el cual ésta información reside no se supervisa típicamente y el contenido es prácticamente invisible a los interesados. La búsqueda es generalmente una búsqueda indexada.” [MU023]

Según Magnus Stensmo y Mikael Thorson, "La Gestión de la Información no Estructurada consiste en las herramientas y métodos necesarios para almacenar, acceder y recuperar, navegar y generar conocimiento primordialmente de la información basada en texto", definición que se adopta para el presente estudio. [ST024]

7. PYME

El Congreso de Colombia decreta en la ley 905 de 2004 en el Capítulo I, artículo 2 las siguientes definiciones: “Artículo 2º. Definiciones. Para todos los efectos, se entiende por micro incluidas las Famiempresas pequeña y mediana empresa, toda unidad de explotación económica, realizada por persona natural o jurídica, en actividades empresariales, agropecuarias, industriales, comerciales o de servicios, rurales o urbanos, que responda a dos (2) de los siguientes parámetros:

1. Mediana empresa:

a) Planta de personal entre cincuenta y uno (51) y doscientos (200) trabajadores, o

b) Activos totales por valor entre cinco mil uno (5.001) a treinta mil (30.000) salarios mínimos mensuales legales vigentes.

2. Pequeña empresa:

a) Planta de personal entre once (11) y cincuenta (50) trabajadores, o

b) Activos totales por valor entre quinientos uno (501) y menos de cinco mil (5.000) salarios mínimos mensuales legales vigentes” [CO025]

8. SITUACION ACTUAL

Según la Superintendencia de Sociedades y Confecamaras, las Pymes representan el 27% de las empresas en nuestro país, 22% pequeñas y 5% medianas distribuidas en todos los sectores de la economía. Las Pymes son una de las principales fuerzas generadoras de empleo, junto a las microempresas acumulan el 75% del total del empleo y, en contraste a esta realidad, esta proporción no se ve reflejada en sus ventas, donde la gran empresa es la que domina en los ingresos. [DA026]

Estas comparaciones con las grandes empresas se hacen debido a que “permitirá tener un panorama más amplio de evaluación, considerando que este tipo de empresas suponen estructuras y organizaciones eficientes que les permiten conservar sus activos productivos en condiciones competitivas bajo un mercado cambiante y cada vez más exigente” [DA026]. Aquí es importante resaltar que esas estructuras y organizaciones eficientes se alcanzan, entre otras cosas, apoyando los procesos con tecnología, lo cual implica puestos de

trabajo más calificados. Además, estas tecnologías no son asequibles en muchos casos a la pequeña o mediana empresa debido a su costo y, la mayoría de las veces no es creada en nuestro país.

En contraste con las empresas grandes, las Pymes no tienen las mismas facilidades de acceso para el financiamiento. Esto se debe a múltiples factores, tales como las tasas de interés, los plazos impuestos para cancelar los créditos, los requisitos de las entidades bancarias, las garantías y codeudores exigidos. Estos factores han influido de tal manera que muchas prefieren financiarse a partir del capital propio. En el caso concreto de las pymes desarrolladoras de software, un factor adicional es el hecho de que las instituciones financieras no aceptan intangibles como garantías.

Además de verse afectadas por la situación económica, también les afecta la situación social actual, específicamente el orden público, la violencia y la inseguridad. En algunas ocasiones las Pyme se ven obligadas a cerrar o reparar alguna sucursal o se ven afectados los canales de distribución dificultando más aún el acceso al mercado interno, el cual también se ve afectado por el contrabando y la competencia desleal.

El interés de la Pyme por las exportaciones es muy bajo debido a que no se percibe como una oportunidad clara para su desarrollo. Esto se debe a que no se tiene conocimiento de los mercados externos que puedan demandar la diversa oferta de esta clase de empresas y por los trámites con sus respectivos costos que una actividad como esta conlleva.

Los encargados de la gestión de la Pyme sienten que la calidad del recurso humano no es la indicada para desempeñar las actividades requeridas. Esta

inconformidad se percibe a través de todas las áreas que componen la empresa. Adicionalmente, un gran porcentaje de la Pyme no cuenta con los servicios de consultoría y capacitación. [DA026] y [GE027]

9. DISEÑO DE LA ENCUESTA

9.1 JUSTIFICACIÓN

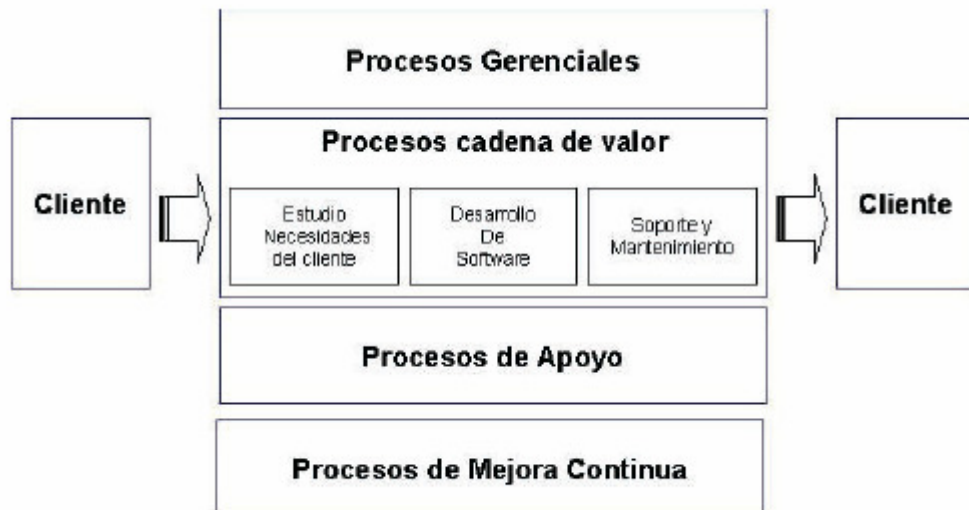
Es evidente que en la mayoría de las organizaciones, ya sean pequeñas o grandes, o dedicadas a cualquier sector económico, se utiliza y se genera más información no estructurada que estructurada. Es tanto así, que recientes estudios estiman que el 80% de la información que generan las empresas es no estructurada. [IN027] y [FE006]

A raíz del gran impulso que ha tenido la información no estructurada en el ámbito mundial, se pretende conocer cual es el manejo actual que se le ha venido dando a dicha información en las Pymes Desarrolladoras de Medellín por medio de entrevistas en el campo tecnológico.

9.2 DISEÑO MUESTRAL

Debido a que el alcance de la población objeto de estudio está limitado a la pequeña y mediana industria tecnológica, concretamente Desarrollo de Software en la ciudad de Medellín, el estudio incluye la mayoría de estas organizaciones.

Las organizaciones que se abarcan dentro del grupo denominado Desarrollo de Software son aquellas que en su proceso de negocio se dedican a algunas de las siguientes actividades: Ingeniería de Requisitos, Diseño y Arquitectura de Software, Desarrollo, Pruebas y Mantenimiento.



Gráfica 1. Mapa de Procesos de una Empresa Desarrolladora de Software. [LO036]

Según la Cámara de Comercio de Medellín para Antioquia, las empresas pequeñas y medianas dedicadas al desarrollo de software establecidas en la ciudad de Medellín son 41. [CC030] El número de empresas entrevistadas fue 21, las cuales representan el 51% de la población objeto de estudio.

9.3 ELABORACIÓN DEL CUESTIONARIO

En un principio se intentó realizar una encuesta tipo cuestionario directo. Esta encuesta llevaba las siguientes preguntas:

En el contexto empresarial

1. ¿Conoce el concepto de Información no Estructurada?
Si ___ No ___
2. ¿Por cual medio se enteró?
 - a. Academia

- b. Internet
 - c. Compañero laboral
 - d. Otro
- Cual _____

3. ¿Utiliza el concepto de Información no Estructurada en la empresa?
Si __ No __

4. ¿Cuales de las siguientes estructuras utiliza?

- a. Documentos de texto
- b. Hojas de Cálculo
- c. Presentación en diapositivas
- d. Portal Web u otros documentos HTML
- e. Imágenes
- f. Video
- g. Audio
- h. Correos Electrónicos

5. ¿Qué herramientas informáticas utiliza para consultar, modificar y almacenar Información No estructurada?

- a. Herramientas office
- b. Navegador
- c. Visor imágenes
- d. Reproductores de audio y video
- e. Manejador de Correo Electrónico

6. ¿Gestiona la Información no Estructurada?

Si __ No __

6.1 ¿Utiliza técnicas de almacenamiento?

Si ___ No ___

6.2 ¿Utiliza técnicas de búsqueda?

Si ___ No ___

6.3 ¿Utiliza alguna herramienta informática para gestionar la Información no Estructurada?

Si ___ No ___

Este tipo de cuestionario puede generar sesgos en la contestación de las personas. Por esto se realizó otro tipo de cuestionario, denominado “encuesta estructurada”. Este nuevo formato intenta llevar al entrevistado a responder las preguntas de una forma más imparcial y más objetiva; muestra cuáles son las reales falencias de la empresa y el nivel de conocimiento que se tiene acerca de la información no estructurada. Así mismo, se puede identificar si las personas que están respondiendo la encuesta están inventando.

La encuesta ha sido organizada teniendo en cuenta el orden de los objetivos y el orden de la entrevista, con el fin de que cada pregunta corresponda con cada uno de los objetivos del estudio y de evitar el sesgo en la información. El siguiente es el formato tipo “encuesta estructurada” según el orden de los objetivos que se utilizó:

1.

1.1 ¿Que conoce acerca de la Información no estructurada?

1.2 ¿Qué diferencias hay entre la Información no estructurada y la estructurada?

1.3 ¿Qué porcentaje de información que maneja es no estructurada?

2.

2.1 ¿Utiliza bases de datos?

2.2 ¿En la escala de 1 (débil) a 10 (fuerte), en que medida las decisiones están basadas en la información de las bases de datos?

2.3 Si no tiene bases de datos, ¿Cómo almacena la información?

3.

3.1 ¿Utilizan correo electrónico? ¿Qué áreas? ¿Qué cargos? ¿Es interno?

3.2 ¿Utilizan videos o animaciones dentro o fuera de la empresa para explicar procesos o procedimientos?

3.3 ¿La empresa tiene un portal Web? ¿Para qué? ¿Hacen mantenimiento?

3.4 ¿Se utilizan imágenes para soportar algún proceso? ¿Cuáles procesos? ¿Qué método? (Correo o TIC's)

3.5 ¿Manejan la información en que tipo de documentos? puede ser electrónico o tangible

3.6 ¿Se utiliza audio para soportar algún proceso? ¿Cuáles procesos? ¿Qué método? (Correo o TIC's)

4.

4.1 ¿Qué herramientas informáticas utiliza para el apoyo de los procesos en la empresa?

4.2 ¿De esas herramientas cuales generan información que agreguen valor a la empresa?

5.

5.1 ¿Cómo almacena sus documentos y demás contenidos?

5.2 ¿Cómo busca sus documentos y demás contenidos?

5.3 ¿Es posible consultar el contenido de los documentos y de que forma se hace en la empresa?

5.4 ¿Tiene forma de observar las relaciones que hay entre el correo electrónico, las imágenes, el portal Web, texto, video, audio, papeles tangibles?

6.

6.1 De acuerdo a 5.4, ¿Se toman decisiones basándose en estas relaciones?

9.4 PERSONAS ENTREVISTADAS

Las personas seleccionadas fueron aquellas que tienen un contacto directo con la gestión, extracción, consulta, desarrollo y mantenimiento de la información en las diferentes pymes. Claro que, debido a los diferentes cargos de las empresas, en muchas ocasiones no existe una persona que esté involucrada directamente con cada uno de los conceptos anteriormente mencionados. El número de personas entrevistadas fue el mismo número de la muestra (21 personas), es decir, una persona por empresa. Los cargos de estas personas son 15 Gerentes Generales y un Analista de Soporte, un Analista de Plataforma, un Director de Gestión, un Coordinador Help Desk, un Analista Programador y un Administrador de Plataforma Tecnológica.

10. PROCEDIMIENTO DE ADMINISTRACIÓN

Este estudio se ha conducido principalmente por medio de las técnicas usuales para obtener datos cualitativos de la población anteriormente definida, a saber: entrevistas personalizadas con un cuestionario previamente desarrollado. Cuando no fue posible esta vía de comunicación directa se optó por entrevista a través de llamadas telefónicas. En última instancia se recurrió al correo electrónico.

11. ANÁLISIS

11.1 Respuestas más relevantes sin variabilidad

A continuación se pueden identificar los componentes que dan guía a las entrevistas, en donde se puede evidenciar según los conocimientos de los entrevistados, que no existe variabilidad. Esto quiere decir que el 100% de las personas contestaron en una pregunta la misma respuesta. No obstante, por medio de este hecho particular se puede demostrar la presencia de la información no estructurada en las pequeñas y medianas empresas de desarrollo de software.

11.1.1 *¿Utilizan correo electrónico?*

De acuerdo con los datos de la encuesta se puede decir que el 100% de las empresas utilizan correo electrónico. Esto evidencia la gran importancia que esta herramienta tiene para las empresas, tanto para la comunicación entre sus empleados (internamente) como para los clientes, proveedores, usuarios (externamente).

También se observa, que es de vital importancia el uso y manejo del correo electrónico en todas las áreas. Estas áreas involucran todos los departamentos administrativos (contable, facturación, etc.) como los departamentos de desarrollo. (Levantamiento de requisitos, análisis, pruebas etc.)

12. RESUMEN

Este resumen pretende enunciar de forma breve, las respuestas que se presentaron después de realizada la investigación. Se mostrarán las respuestas tabuladas de las preguntas cuantitativas y se enunciarán los casos más importantes de las preguntas cualitativas.

¿Utilizan correo electrónico?	Si	100%
	No	0%
¿Qué áreas?	Todas las áreas	100%
¿Qué cargos?	Todos los cargos	100%
¿Utilizan videos o animaciones dentro o fuera de la empresa para explicar procesos o procedimientos?	Si	62%
	No	38%
¿La empresa tiene un portal Web?	Si	100%
	No	0%
¿Hacen mantenimiento al Portal Web?	Si	100%
	No	0%
¿Cada cuánto?	Cada mes	38%
	Cada 6 meses	10%
	Cada Año	19%
	Indefinido	33%
¿Se utilizan imágenes o diagramas para soportar algún proceso?	Si	76%
	No	24%
¿Utiliza algún tipo de sistema para procesar las imágenes ó diagramas? (viene de la pregunta anterior)	Gestión Documental	25%
	No Tiene	75%
¿Manejan la información en que tipo de documentos: electrónicos o tangibles?	Ambos	100%
	Solo Electrónico	0%
	Solo Tangible	0%
¿Se utiliza audio para soportar algún proceso?	Si	48%
	No	52%
¿Qué herramientas informáticas utiliza	Entornos de interfaces de desarrollo.	

para el apoyo de los procesos en la empresa?	Herramientas de modelado de software.	
	Sistemas operativos.	
	CRM's (Customer Relationship Management).	
	Motores de bases de datos.	
	Contabilidad, facturación, gestión de proyectos, exploradores de Internet, aplicaciones para copias de seguridad.	
	Herramientas desarrolladas para su manejo interno como aplicaciones para el manejo de hojas de vidas.	
¿De esas herramientas cuales generan información que agreguen valor a la empresa?	El 43% de las empresas considera que todas sus herramientas informáticas generan información que agrega valor a la empresa.	
	El 19%, sin incluir las que respondieron que todas, advierte que las aplicaciones que generan indicadores, consolidaciones y reportes son las que permiten agregar valor porque facilitan el seguimiento de los procesos.	
	El 24% de las empresas respondieron que las aplicaciones de gestión de proyectos son las más importantes.	
	Un 5% de las empresas aclara que las herramientas informáticas no son las que generan información que agregue valor.	
¿Cómo almacena sus documentos y demás contenidos?	Sólo el 19% de las empresas pueden referenciar los contenidos a parte de los directorios a través del sistema de gestión documental.	
	El 81% lo realizan a través de un tema, fecha (especialmente año), proyecto o clientes.	
¿Cómo busca sus documentos y demás contenidos?	El 57% de las empresas respondieron explícitamente que las búsquedas se realizan por medio del buscador y el explorador del sistema operativo Microsoft Windows.	
	El 43% restante manejan un Sistema de Gestión Documental o alguna aplicación creada por ellos mismos.	
¿Tiene forma de observar las relaciones que hay entre el correo electrónico, las imágenes, el portal Web, texto, video, audio, papeles tangibles?	Es posible	33%
	No es posible	57%
	No responde	10%
¿Utiliza bases de datos?	Si	48%

	No	52%
¿En que medida se basan las decisiones en las bases de datos?	0-25%	10%
	26-50%	14%
	51-75%	19%
	76-100%	57%
¿Qué conoce acerca de la Información no estructurada?	Nada	71%
	Poco	19%
	Mucho	10%
¿Qué diferencias hay entre la Información no estructurada y la estructurada?	La información estructurada es aquella que se puede indexar fácilmente, se tiene documentada y almacenada en cualquier tipo de medio, que proporciona organización y eficiencia, facilidad de búsqueda, facilidad de respaldos.	
	La información no estructurada es aquella que no se puede ubicar ni interrelacionar fácilmente, que está en el conocimiento o experiencias de las personas que hacen parte de la empresa y con la cual básicamente se toman decisiones más relevantes. Además tiene una forma diferente a la estructurada.	
¿Qué porcentaje de información que maneja es no estructurada?	0-25%	24%
	26-50%	14%
	51-75%	19%
	76-100%	19%
	No sabe	24%

Tabla 1. Resumen Encuesta

El 100% de los encuestados, sin importar su cargo dentro de la compañía, utilizan correo electrónico. Involucrando de antemano, el gerente, subgerente, jefes de departamento, auxiliares, desarrolladores, analistas y secretarias, como también a todo el personal administrativo.

Como se pudo observar, las empresas de desarrollo de software manejan correo electrónico generando un gran volumen de información no estructurada. Entre los comentarios que se presentaron al realizar las entrevistas, se destaca que muchas de estas personas no tienen un control preciso de todos los correos que manejan porque no se gestionan. Inclusive, han tenido pérdida de información.

11.1.2 ¿La empresa tiene un portal Web?

El 100% de las empresas de desarrollo de software tienen habilitado en estos momentos un portal Web. Existen dos tipos de portales Web identificados en esta población: Uno es el portal interno para los empleados y otro un portal externo para los clientes en general. El 33% (7 Empresas) tienen ambos portales, mientras que el 66% (14 Empresas) maneja un portal externo únicamente. Dentro de los usos que cada compañía le da a su portal podemos mencionar los siguientes:

- Portal interno para procedimientos y capacitaciones y un portal externo dónde se explican los productos y servicios de la empresa.
- Información empresarial en donde se explica que hace la compañía., información de los clientes, proyectos más importantes o casos de éxito, recepción de hojas de vida y explicación de los productos y servicios.
- Posicionamiento de marca, apoyo comercial y relaciones con los clientes. Además, información acerca del proceso de ingeniería con sus respectivas certificaciones y actualización de la información de la empresa.

- Acceso a los demos y prototipos de las aplicaciones desarrolladas para los clientes.
- Simplemente para hacer presencia en Internet.

11.1.3 ¿En qué tipo de documentos manejan la información?

En esta pregunta se comprobó que el 100% utilizan las herramientas ofimáticas en donde se incluyen procesadores o visores de textos, hojas de cálculo y presentaciones.

Es necesario aclarar que el 48% (10 Empresas) de las organizaciones afirma que las herramientas ofimáticas generan información que le agregan valor a la empresa, ya que en algunos casos se puede encontrar información contable o plantillas para documentación del software.

11.1.4 ¿Utiliza bases de datos?

Las empresas desarrolladoras de software apoyan por lo menos un proceso en las bases de datos relacionales, comúnmente la información concerniente a los proyectos de software y de los clientes.

A partir de esa información se pueden generar reportes, tabulaciones, investigación de mercados, rendimientos e indicadores que apoyan la toma de decisiones.

11.2 Respuestas más relevantes con variabilidad

A continuación se pueden observar los componentes de las entrevistas que si presentaron variabilidad, en donde se clasifican los demás contenidos en los que se puede encontrar información no estructurada y su respectiva relación con el apoyo a la toma de decisiones.

11.2.1 ¿Cada cuánto se hace mantenimiento al portal Web?

Esta pregunta pretendía identificar la coherencia entre la información no estructurada que se encuentra en el portal Web y la actualidad de la empresa. Los datos que se evidenciaron se muestran en la figura 1.

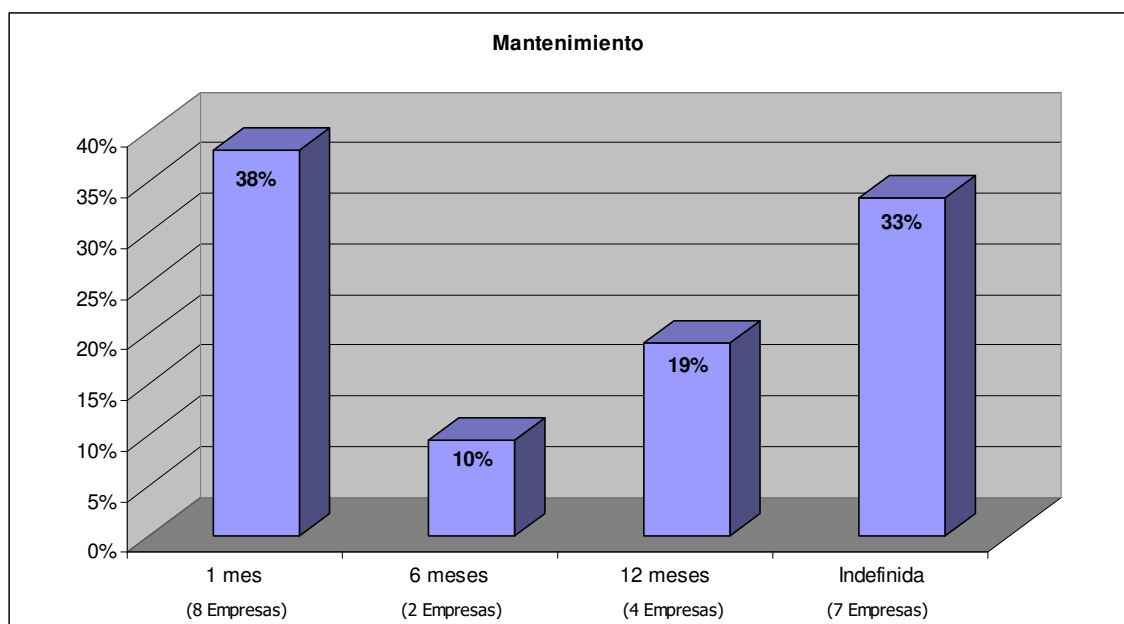


Figura 1. Periodicidad del mantenimiento del portal web.

La periodicidad indefinida varía en el rango de un mes a un año ya que no se pudo establecer una cifra exacta con el entrevistado, esto quiere decir que generalmente el mantenimiento en estos casos se realiza por demanda.

11.2.2 *¿Utilizan videos o animaciones dentro o fuera de la empresa para explicar procesos o procedimientos?*

El 62% (13 Empresas) de las empresas entrevistadas respondió que sí utilizan videos o animaciones para explicar algún proceso o procedimiento perteneciente a la compañía. De esta parte el 46% (6 Empresas) afirma que se utiliza con el fin de capacitar a sus empleados en el manejo de las herramientas y métodos comunes en el proceso de desarrollo de software. El 39% (5 Empresas) afirma que se utiliza para la presentación de sus productos a los clientes. Finalmente, en una menor cantidad, se utiliza con el fin de promocionar la empresa.

El 38% (8 Empresas) restante respondió que no utiliza ninguna animación ni video.

11.2.3 *¿Se utilizan imágenes o diagramas para soportar algún proceso?*

El 76% (16 Empresas) de las empresas entrevistadas respondió que sí utilizan imágenes o diagramas en algún proceso o procedimiento perteneciente a la compañía. Todas las empresas utilizan un método de modelación en su proceso de desarrollo de software. Pero también se genera información no estructurada a partir de las siguientes variables:

- Sistema de gestión y calidad de la empresa.
- Imágenes de las aplicaciones para capacitación de los clientes
- Cronograma de actividades y gestión de proyectos
- Imágenes de las aplicaciones para comunicaciones internas

- Modelación y documentación de procesos que ayudan a las diferentes certificaciones, como también la posibilidad de digitalizar documentos físicos debido a su importancia para el negocio como: facturas, hojas de vida, visitas, contratos tanto de clientes como de empleados, quejas y reclamos, correspondencia entrante y saliente.

El 24% (5 Empresas) restante respondió que no utilizan imágenes en sus procesos.

Sólo un 25% (4 Empresas) tiene herramientas apropiadas para la gestión de las imágenes como un sistema de gestión documental. Esto implica que hay un 75% (12 Empresas) que simplemente tienen la información (imágenes en este caso) almacenadas sin tener algún tipo de método que les ayude a gestionar este tipo de información.

11.2.4 ¿Se utiliza audio para soportar algún proceso?

El 48% (10 Empresas) de las empresas entrevistadas respondió que si utilizan audio en algún proceso o procedimiento perteneciente a la compañía, en especial aquellos procesos donde hay un contacto con los clientes y en los cuales se quiere dejar una constancia de lo que se expresó y se arregló entre las partes. Es decir, esto tiene que ver con la definición del dominio del producto a desarrollar, los contratos y el levantamiento de requisitos. También se utiliza el audio como parte de otro documento para explicar algún proceso o dar una capacitación.

El 52% (11 empresas) restante respondió que no utiliza audio.

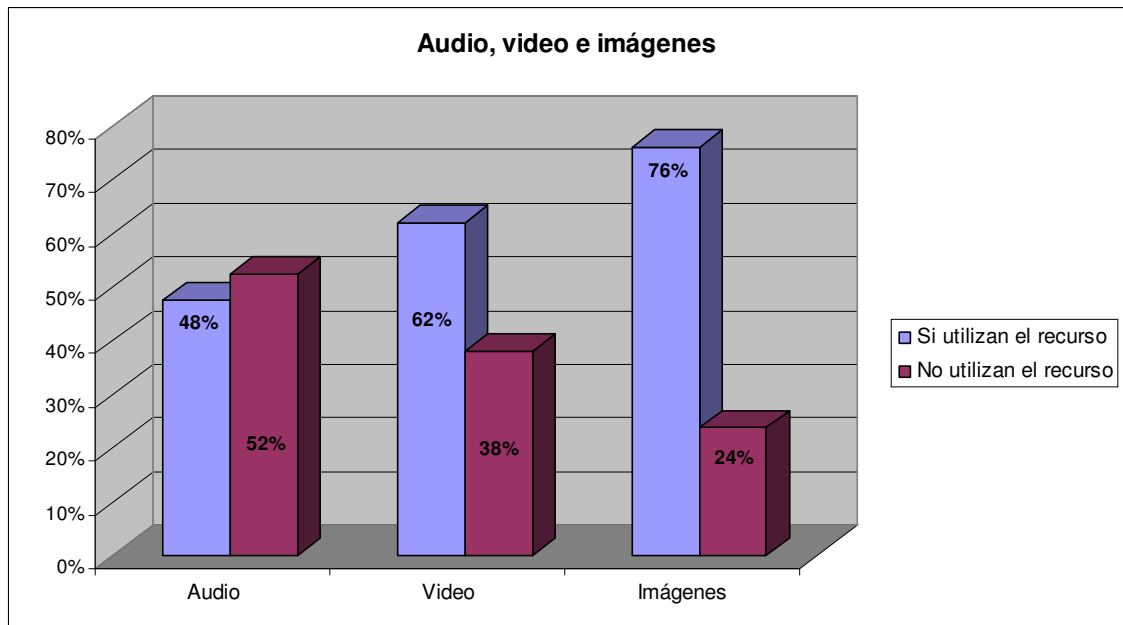


Figura 2. Resumen porcentual de Audio, video e imágenes.

Todas las empresas que han utilizado audio (10 Empresas) también manejan video e imágenes dentro de sus procesos o procedimientos.

Como el audio hace parte del video el 23% (3 Empresas) no utiliza este contenido. Según las respuestas encontradas no se hace necesario ya que, a partir del video, se puede extraer el audio que les sirva para un proceso o procedimiento determinado. (Ver Figura 3)

Mientras tanto, el 100% siguen utilizando las imágenes en el negocio como se puede evidenciar en la figura 3.

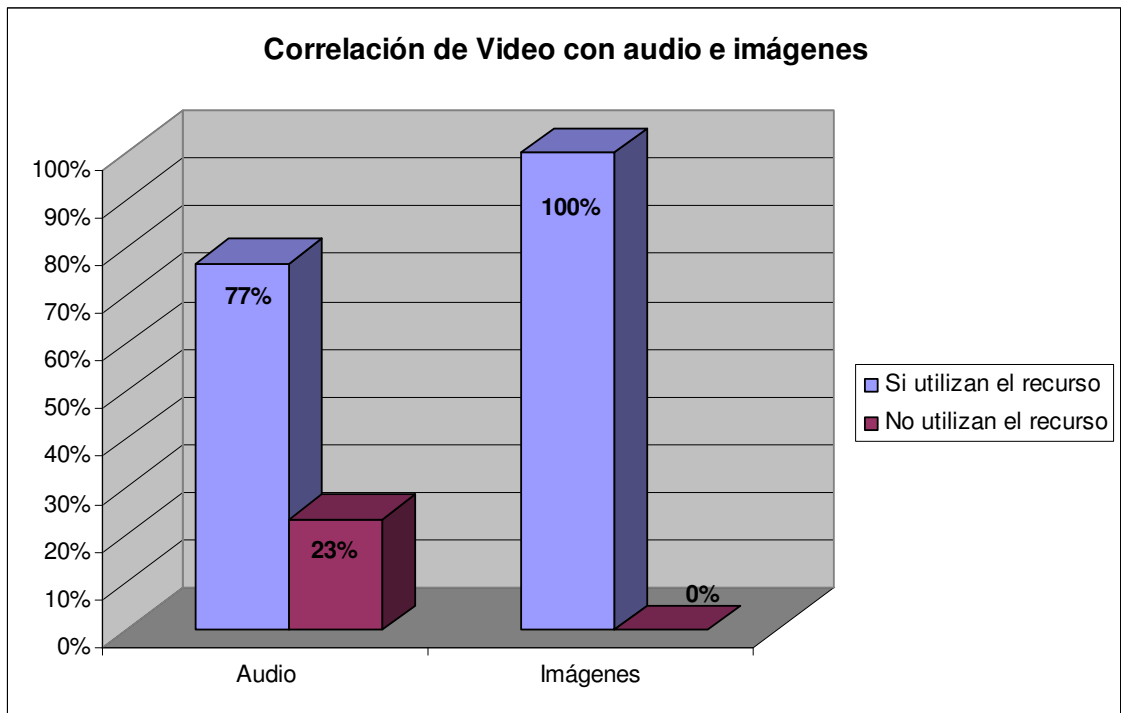


Figura 3. Correlación de Video con audio e imágenes.

Este es el caso más particular ya que, según los datos, las imágenes son las más utilizadas por las empresas. Pero esto no quiere decir que exista una directa relación con la utilización de video y audio ya que en ninguno de los dos se presenta en un 100%. Inclusive, solo el 67% (10 empresas) utilizan audio sabiendo de antemano que manejan imágenes. Mientras que en un 87% (13 Empresas) se manipula el video. Por esto, las imágenes son parte fundamental de este tipo de negocio ya que sin importar cuál es el contenido global de las empresas, éstas se hacen imprescindibles para la organización, tanto para el desarrollo de sus productos como para la gestión de los mismos.

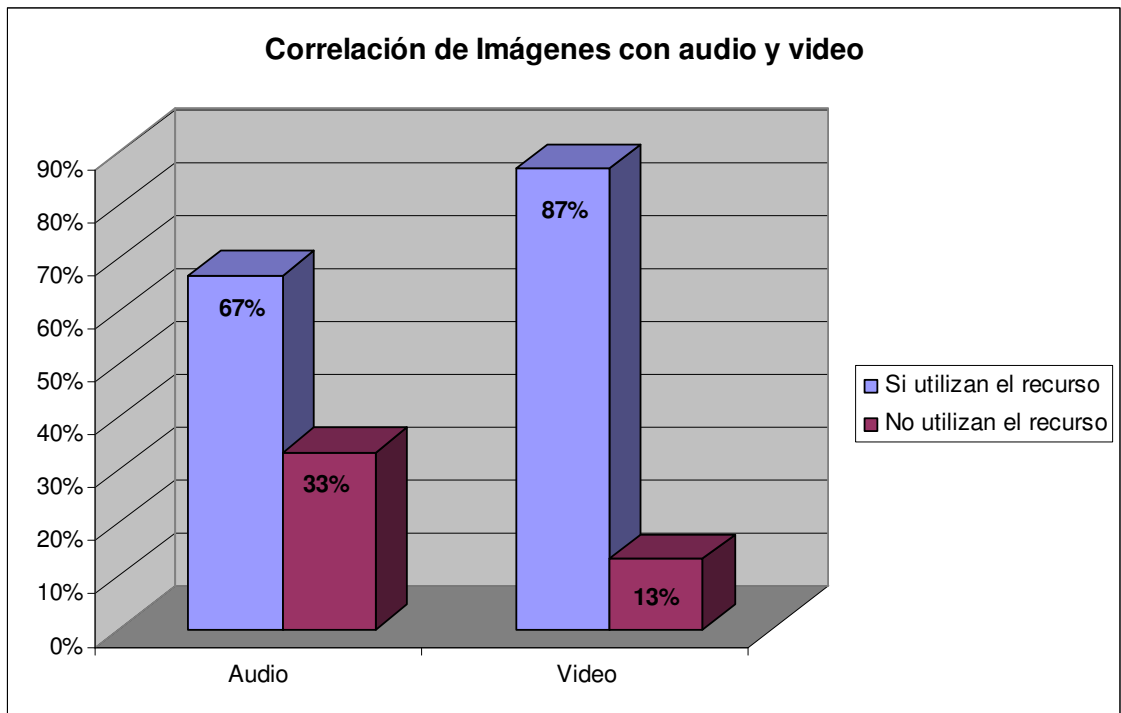


Figura 4. Correlación de Imágenes con audio y video.

Algunas empresas manifestaron pretender utilizar el audio y/o el video en futuras ocasiones con el fin de tener un respaldo, especialmente en certificaciones.

11.2.5 ¿Manejan la información en qué tipo de documentos?

El 100% de las empresas manejan sus documentos de forma electrónica y tangible. Sin embargo, se encontró una tendencia a eliminar toda la documentación física. En la mayoría de casos, las organizaciones tienen aún un archivo físico en donde los documentos que se determinen importantes, se digitalizan. Se pudo identificar en tres casos que todo documento físico tenía su respectiva imagen digital mientras que en otro caso se pudo constatar que todo archivo físico tenía su imagen digital y todo documento digital tenía su presencia física en el archivo.

En todo caso, las compañías mantienen de forma organizada su archivo físico, pero esto no hace que sea fácil su gestión. De hecho, como las búsquedas son manuales conllevan mucho más tiempo y la información es difícil de relacionar.

11.2.6 *¿Qué herramientas informáticas utiliza para el apoyo de los procesos en la empresa?*

Dentro de las principales herramientas que generan información, ya sea estructurada o no estructurada, encontramos las siguientes:

- Entornos de interfaces de desarrollo.
- Herramientas de modelado de software.
- Sistemas operativos.
- CRM's (Customer Relationship Management).
- Motores de bases de datos.
- Contabilidad, facturación, gestión de proyectos, exploradores de Internet, aplicaciones para copias de seguridad.
- Herramientas desarrolladas para su manejo interno como aplicaciones para el manejo de hojas de vidas.

Es importante resaltar que solo el 19% (4 Empresas) utilizan un sistema de gestión documental ya que esta herramienta permite manejar cierta información no estructurada, en su mayoría textos, de manera indexada.

11.2.7 ¿De esas herramientas cuáles generan información que agreguen valor a la empresa?

A pesar de que el proceso central de la población objeto de estudio es el de desarrollo de software, el 43% (9 Empresas) considera que todas sus herramientas informáticas generan información que agrega valor a la empresa. Además, afirman que en caso de que alguna de sus herramientas no agregue valor a la empresa sería eliminada. Un 5% (1 Empresa) aclara que las herramientas informáticas no son las que generan información que agregue valor sino que son las personas las encargadas de esta función luego de agregarle criterio a los datos arrojados por las aplicaciones.

El 62% (13 Empresas), incluyendo las empresas que respondieron que todas las herramientas generan valor, indicaron que las aplicaciones de desarrollo también lo hacen. Así mismo, un 19% (4 Empresas), sin incluir las que respondieron que todas, advierte que las aplicaciones que generan indicadores, consolidaciones y reportes son las que permiten agregar valor porque facilitan el seguimiento de los procesos. Por último, el 24% (5 Empresas) respondieron que las aplicaciones de gestión de proyectos son las más importantes.

Un caso específico que vale la pena resaltar es que en una de las empresas se logró sacar un plan de contingencia para solucionar una necesidad concreta en cuestión de pocos días, solo por tener de cierta forma, una estructura de la información que típicamente en otras empresas es no estructurada. Dicha empresa solía tener en promedio 5.000 llamadas por mes por parte de todos los clientes. Al año siguiente en los meses de febrero y marzo, ésta información se incrementó en un 30% generando congestiones en el sistema. Al revisar las llamadas a través del sistema se dieron cuenta que la mayoría de las llamadas eran causadas por un mismo factor. Al percatarse del inconveniente, llevaron a

cabo un plan de contingencia que consistía en realizar tres conferencias para todos los clientes, aclarando en ellas cuál era el inconveniente que se estaba presentando. La importancia de tener en las manos este tipo de información, radica en que sólo con un reporte hecho por el sistema se logró detectar y reducir los problemas que se presentaban. No necesitaron contratar más personal para el área de atención al cliente, manteniendo los índices de servicio en un 90%. Claro que en este campo no sólo se tiene un valor agregado. Existe un módulo de quejas y reclamos donde se tiene información precisa de la fecha, el nombre del cliente, la vigencia del contrato, el personal que atendió el servicio, la ubicación del archivo físico con las firmas del visitante y el cliente, más la queja o reclamo en sí. Y todo esto se puede extraer solo generando un reporte en el sistema.

Otro de los casos importantes es el del uso de una herramienta denominada Twiki, la cual es un Software de código abierto creado y mantenido por un grupo de muchas personas encabezado por Peter Thoeny. [TW037] Twiki es un sistema que se utiliza como gestor de documentos en dónde se documenta todo el software con el manejo de versiones. Con esta herramienta es posible desarrollar aplicaciones web sin necesidad de insertar código. Así mismo, provee un control de acceso sofisticado opcional lo que hace más eficiente la gestión de la información. Para este caso se observó de forma directa la posibilidad de realizar búsquedas entre los diferentes proyectos y procesos, accesos a los diferentes correos electrónicos, anexos de diferentes archivos (video, audio, imágenes, textos, etc.) como también una estructura lógica del contenido formado por metadatos. Es importante mencionar que con esta herramienta es posible consultar en las bases de datos, ordenar tablas y manejar hojas de cálculo. También es posible hacer un seguimiento detallado de los procesos y los cambios en la documentación, que son visibles instantáneamente para todos los empleados de las empresas debido a la filosofía Wiki, la cual se refiere a un “sitio web colaborativo que puede ser

editado por varios usuarios. Los usuarios de una wiki pueden así crear, modificar, borrar el contenido de una página web, de forma interactiva, fácil y rápida; dichas facilidades hacen de la wiki una herramienta efectiva para la escritura colaborativa.” [WI031] Fue tan evidente el uso de la información no estructurada que se constató, a través de una búsqueda, diferentes relaciones partiendo del contenido de lo que allí estaba almacenado.

11.2.8 *¿Cómo almacena sus documentos y demás contenidos?*

Todas las empresas utilizan una estructura de directorios y subdirectorios para almacenar sus documentos. La clasificación de estos directorios se hace a la vez por alguno de los siguientes factores, a saber: por tema, fecha (especialmente año), proyecto o clientes. Dentro de cada directorio referido a un cliente se encuentran todos los contenidos relacionados a éste, excepto el correo electrónico y los documentos físicos que no se consideren de vital importancia. Sólo el 19% (4 Empresas) pueden referenciar los contenidos a parte de los directorios a través del sistema de gestión documental.

El manejo de versiones también se ha vuelto común no sólo para código fuente, sino además, para la documentación de los proyectos permitiendo a la empresa omitir modificaciones que se hayan hecho erróneamente. Por último, todas las empresas tienen respaldo de los contenidos garantizando perdurabilidad de los mismos al realizar copias de seguridad en medios magnéticos, servidores y cd's.

11.2.9 *¿Cómo busca y consulta sus documentos y demás contenidos?*

Las búsquedas son muy limitadas ya que muy pocas empresas tienen la oportunidad de añadir metadatos o realizar indexaciones a los contenidos por

medio de una aplicación. Esto se debe en gran parte al no tener un sistema de gestión documental que permita llevar a cabo estas tareas. Como todas las empresas se limitan a utilizar la estructura de directorios para almacenar los documentos se les dificulta posteriormente las búsquedas, las cuales se realizan a través de la herramienta proporcionada por el sistema operativo y parametrizadas por las categorías anteriormente mencionadas (tema, fecha, proyecto o clientes). Es tanto así, que el 57% (12 Empresas) respondieron explícitamente que las búsquedas se realizan por medio del buscador y el explorador del sistema operativo Microsoft Windows.

11.2.10 ¿Tiene forma de observar las relaciones que hay entre el correo electrónico, las imágenes, el portal Web, texto, video, audio, papeles tangibles?

Un 33% (7 Empresas) de los entrevistados respondieron que sí pueden observar relaciones entre los contenidos. Sin embargo, estas relaciones no se pueden obtener a través de una aplicación sino que se realiza por medio de una operación manual. Para ello se utilizan matrices de trazabilidad, proceso que se toma bastante tiempo y que puede resultar tedioso para los empleados que lo realizan. Adicionalmente, dificulta la obtención inmediata de resultados para la organización. También se pueden clasificar todos los contenidos en categorías que dependen de un evento que los relaciona, que a su vez, va de la mano del almacenamiento en estructura de directorios.

Por el contrario, el 57% (12 Empresas) afirma que no es posible relacionar estos contenidos. Inclusive, las respuestas se tornan un tanto más próximas a la relación que hay entre los directorios y sus subdirectorios y no acerca del contenido de esos directorios en sí.

Mientras que un 10% (2 Empresas) no sabe o no responde si tiene forma de observar las relaciones entre los contenidos.

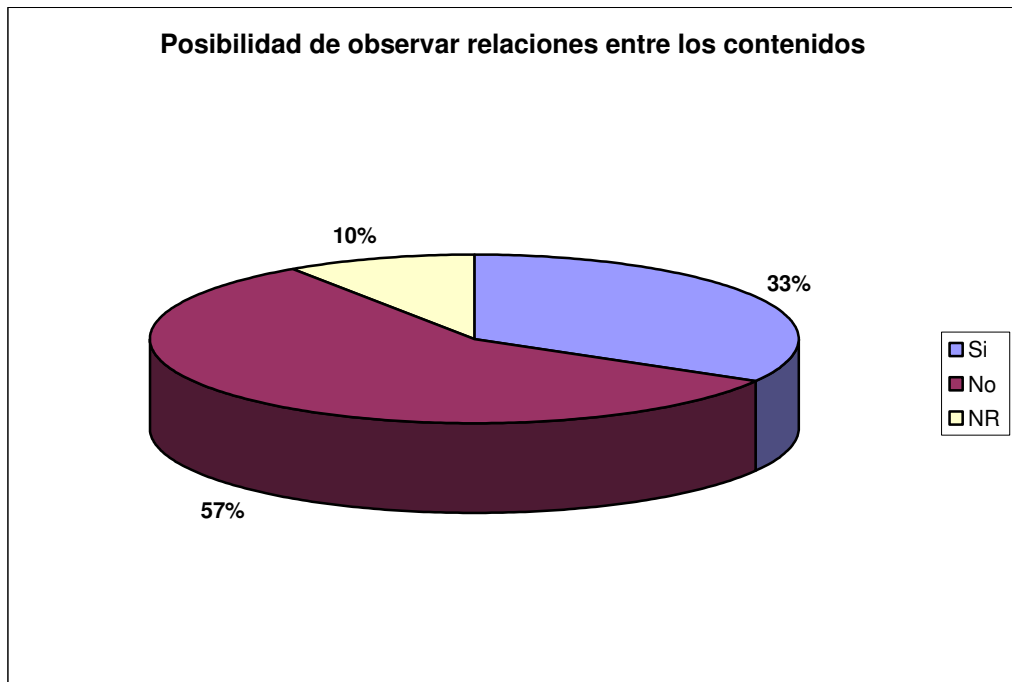


Figura 5. Posibilidad de observar relaciones entre los contenidos

11.2.11 ¿Se toman decisiones basándose en estas relaciones?

Estas relaciones que se establecen manualmente y no de manera automática y que además son de vital importancia son usadas para identificar tendencias y construir planes de contingencia en un instante determinado. Un 43% (3 Empresas) de las empresas que respondió que sí se podían observar las relaciones afirmó que se toman decisiones basándose en estas relaciones. El 57% (4 Empresas) restante respondió que no se tomaban decisiones a partir de esas relaciones.

11.2.12 ¿Qué porcentaje de información está en las bases de datos?

Este tema fue muy difícil de aclarar ya que nunca se pudo verificar a ciencia cierta cual era la cantidad de información que residía en las bases de datos y se trataba más de un dato estimado de la persona entrevistada. Aunque hay

algunos estudios referenciados anteriormente que demuestran que solo el 15% o 20% de la información corporativa es estructurada esta pregunta quería establecer cual era esa proporción y observar si es coherente con esos estudios.

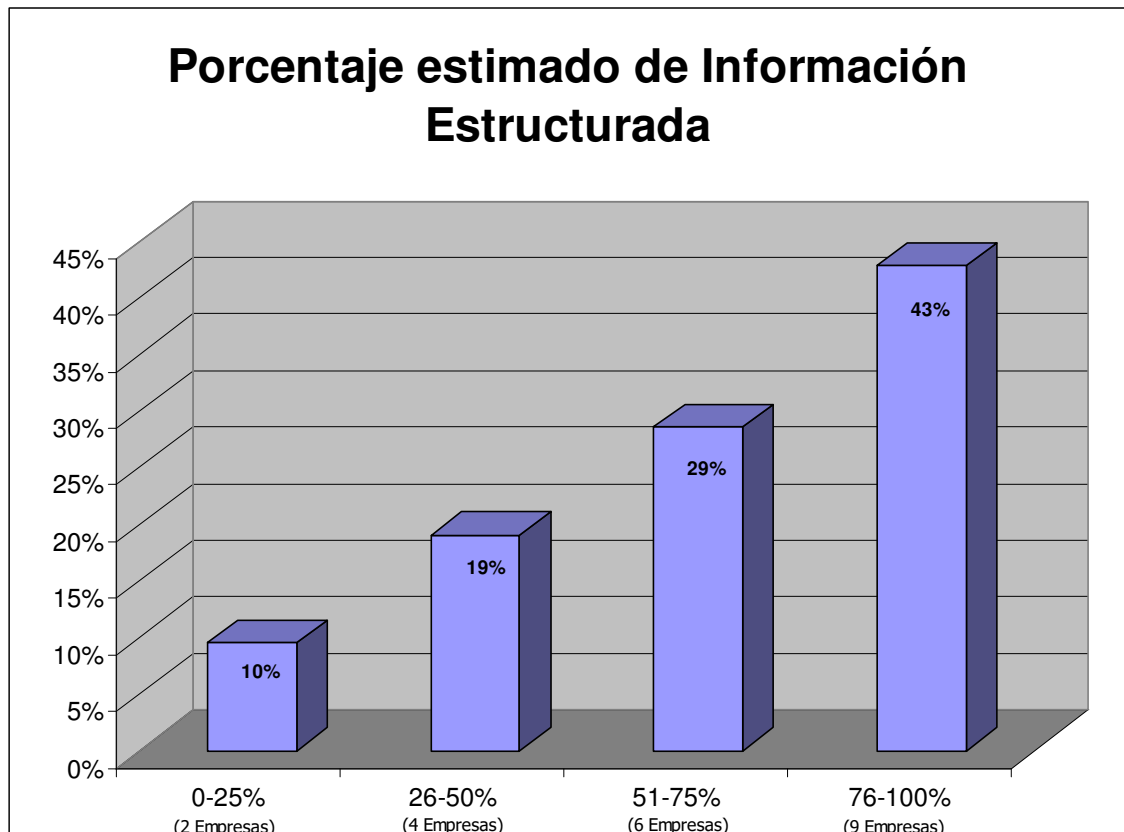


Figura 6. Porcentaje estimado de Información Estructurada

Basándose en la Figura 6 se puede observar que los dos últimos intervalos agrupan el 71% (15 Empresas) de las empresas. Esto implica que este porcentaje de organizaciones aclara tener más del 50% de la información corporativa bajo un esquema rígido del que se puede obtener significado. Tanto así, que el 43% (9 Empresas) afirma tener casi toda la información

estructurada. Por otro lado, solo el 29% (6 Empresas) determina que menos de la mitad de su información corporativa está controlado bajo una estructura.

11.2.13 ¿En qué medida o que porcentaje, la toma de decisiones se apoya en las bases de datos?

En la Figura 7 se puede observar que la toma de decisiones en este sector se basa primordialmente en las bases de datos ya que estas son una fuente de indicadores de rendimiento y generadoras de reportes como lo afirmaron las personas entrevistadas.

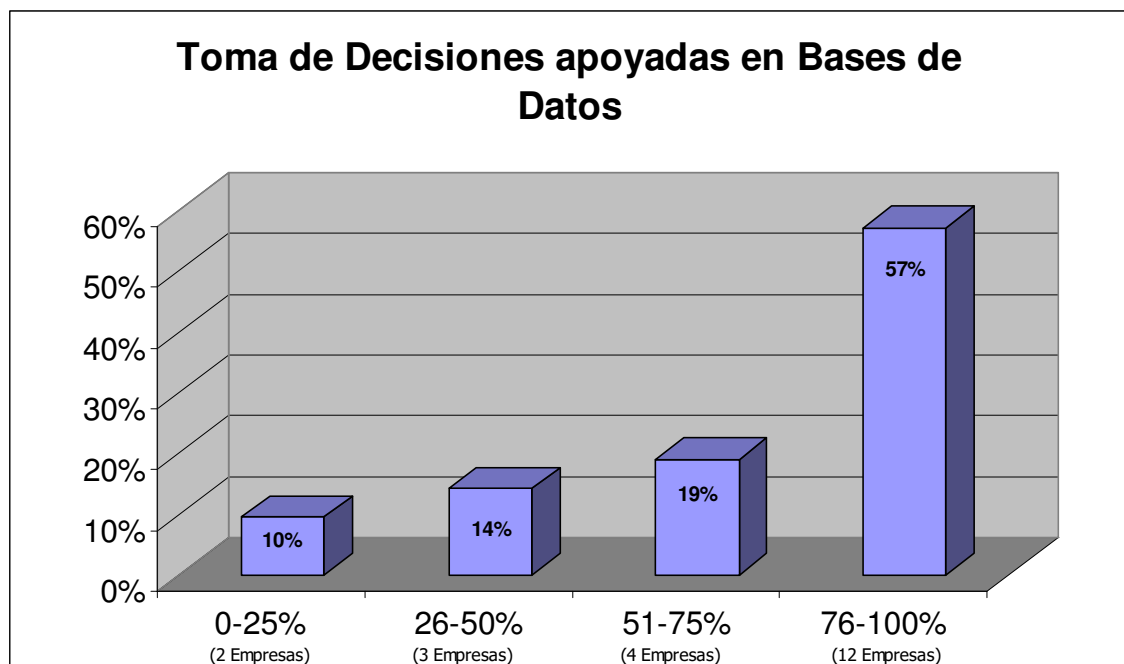


Figura 7. Porcentaje estimado de toma de decisiones apoyadas en Bases de Datos.

El 24% (5 Empresas) asegura que la toma de decisiones se apoya menos del 50% en las bases de datos porque esta información se utiliza en pocos procesos, tales como los administrativos, financieros y de calidad. Se pudo

encontrar un caso atípico en donde la persona entrevistada aseveró no apoyar la toma de decisiones en las bases de datos.

11.2.14 ¿Qué conoce acerca de la Información no estructurada?

El 10% (2 Empresas) entrevistadas, además de citar características y propiedades del tema, estuvieron acertadas en dar una definición teórica del mismo y en estos casos tenían desarrollada una herramienta para tal efecto.

El 19% (4 Empresas) conocen poco acerca de este tema. Algunas de las definiciones recibidas tuvieron que ver con características de la información no estructurada pero no con su definición formal. Entre las respuestas más comunes se encuentran las siguientes:

- Es una información que tiene difícil ubicación e indexación.
- Es información que se puede encontrar en papel, documentos office, conversaciones telefónicas, faxes, etc.

Algoritmos de búsquedas de cosas no previstas. Comentarios de clientes como cosas que se dicen y que se oyen, cosas de la prensa que no se registran en ningún medio.

Los resultados generales muestran que el 71% (15 Empresas) de las organizaciones no conocen nada acerca de la información no estructurada.

De hecho, el término generó en las personas entrevistadas respuestas inapropiadas que inclusive no se asemejan, de ninguna forma, a la respuesta esperada. A su vez, generó reacciones negativas y de mucha incertidumbre para continuar con la entrevista. Y esta es una de las razones por la que la pregunta se encuentra al final de la guía de la encuesta. Inclusive, muchas

personas devolvieron la pregunta buscando tener una idea con respecto a este tema.

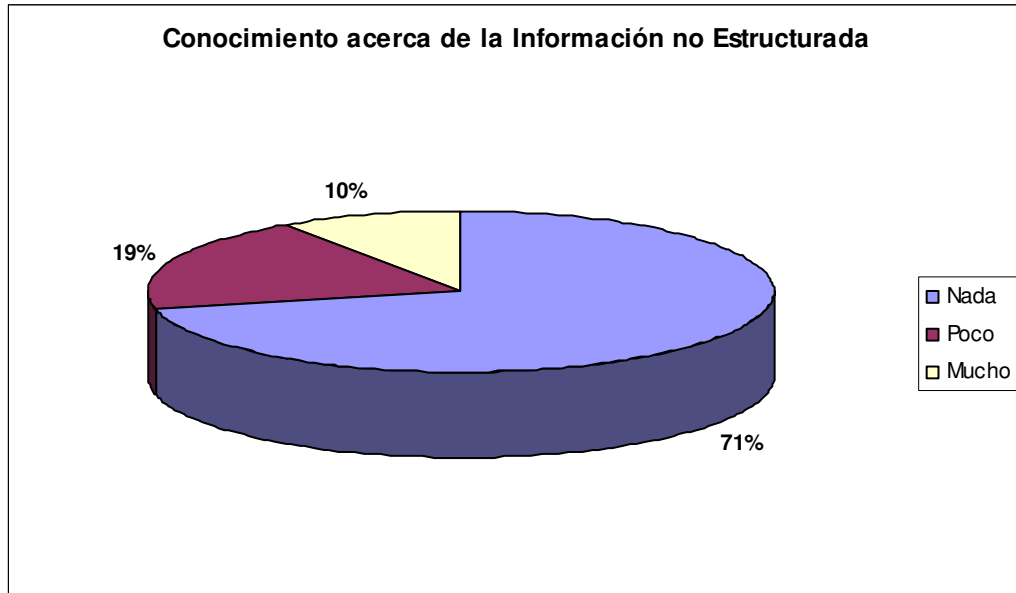


Figura 8. Conocimiento acerca de la Información no Estructurada

11.2.15 ¿Qué diferencias hay entre la Información no estructurada y la estructurada?

Como muy pocas personas conocen acerca del tema se les hace igual de difícil determinar cuáles son las diferencias que hay entre la información estructurada y la no estructurada. A pesar de esto, se pueden rescatar las siguientes respuestas relevantes de un 19% (4 Empresas):

- La información estructurada es aquella que se puede indexar fácilmente, se tiene documentada y almacenada en cualquier tipo de medio, que proporciona organización y eficiencia, facilidad de búsqueda, facilidad de respaldos.

- La información no estructurada es aquella que no se puede ubicar ni interrelacionar fácilmente, que está en el conocimiento o experiencias de las personas que hacen parte de la empresa y con la cual básicamente se toman decisiones más relevantes. Además tiene una forma diferente a la estructurada.

11.2.16 ¿Qué porcentaje de información que maneja es no estructurada?

A diferencia de la pregunta relacionada con la información estructurada en donde se podía observar a simple vista una tendencia mayoritaria cercana al 100%, en esta, se observa que no existe ninguna tendencia hacia un valor en particular. De hecho, no existen muchas diferencias entre las categorías presentadas.

Cabe resaltar, que las personas encuestadas al no conocer acerca de la información no estructurada, respondió en un 24% (5 Empresas) que no saben cuánto porcentaje estimado manejan de este tipo de información.

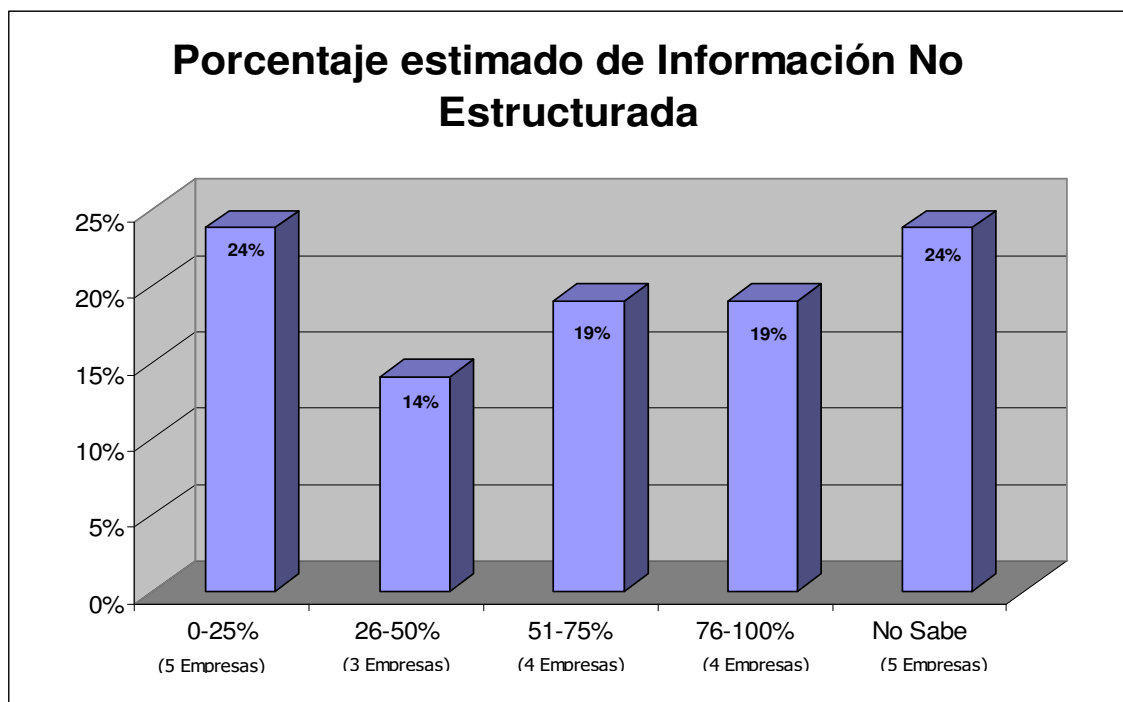


Figura 9. Porcentaje estimado de Información no Estructurada

14. CONCLUSIONES

Según este estudio, las razones por las que las Pymes Desarrolladoras de Software de Medellín no utilizan Sistemas de Información No Estructurados son las siguientes:

- Por la sensación que tienen las empresas de tener toda la información corporativa estructurada.
- Porque en las empresas aun no se han percatado de cuáles son las necesidades específicas que estos sistemas puedan satisfacer.
- Porque se tiene la creencia de que no es tan necesario para este tipo de negocio.
- Porque desconocen el concepto. No han tenido la oportunidad de tener a la mano este conocimiento ni por parte de los empleados, el entorno o la casualidad.

Las siguientes conclusiones se refieren al análisis de los datos:

- Las empresas que tengan un conocimiento medio apropiado de la información no estructurada tienen un mayor campo de acción a la hora de tomar decisiones que generen valor para la empresa. Esto sin incluir que las empresas, teóricamente hablando, no tienen un gran conocimiento acerca de este tipo de información. Sin embargo, son capaces de llevarlo a la práctica y evidenciar gran ventaja con respecto a sus competidores.
- Se evidenció que el correo electrónico es una de las fuentes que más genera información no estructurada en las empresas de este sector, de la misma forma los documentos también generan un alto índice de este tipo de información. En un segundo plano aparecen las imágenes

seguidas por el video y el audio. A pesar de que todas las empresas reciben llamadas telefónicas, realizan contactos con los clientes para definición de requerimientos, reuniones para el seguimiento del producto y contratos, no se respalda esta información en audio. El portal interno ocupa un lugar importante ya que está continuamente en crecimiento y modificación por parte de los empleados con nuevas versiones que incluyen más contenidos. Por último está el portal Web donde debido a sus pocas actualizaciones no genera mucha información no estructurada.

- Los usos más comunes que se le dan al correo electrónico son las comunicaciones con los clientes, entre los mismos empleados o con los proveedores. El audio se usa para el levantamiento de requisitos y la definición de contratos. El video se usa para la capacitación de empleados en algún método o herramienta, para la capacitación de los clientes con un producto nuevo y para la presentación de la empresa. Las imágenes se usan para el proceso de ingeniería, el modelado de procesos en el sistema de calidad y también para capacitaciones y comunicaciones internas. El portal interno se utiliza para documentar los procesos de la empresa, los proyectos, las aplicaciones, herramientas y métodos.
- La parte de la información no estructurada sobre la que se tiene menos control es el correo electrónico. Este fenómeno se da ya que las empresas manipulan el correo como sucesos aislados y además, no hacen parte de ningún sistema de gestión documental.
- La gran mayoría de las empresas no utilizan una herramienta que les permita gestionar la información no estructurada de una manera adecuada. Hasta ahora, pocas empresas recurren a un Sistema de Gestión Documental que facilite esta actividad de una manera más automática. Estas herramientas están enfocadas al manejo de archivos basados en lenguaje natural.

- En las empresas que gestionan la información no estructurada se puede ver un beneficio en la toma de decisiones tan pertinente como si la fuente de la información proviniese de un esquema estructurado.
- En la pequeña y mediana industria del desarrollo del software de la ciudad de Medellín se tiene la sensación de tener toda la información bajo una estructura de la cual se puede obtener todo su significado. Una de las razones por las que se puede explicar este fenómeno es por el desconocimiento de qué es la información no estructurada. Por otro lado, porque la información que se encuentra en las bases de datos relacionales es la de vital importancia para el funcionamiento de este negocio. Por último, debido a que la parte documental de las organizaciones también hace parte de las bases de datos pero no se tiene clara la diferencia con respecto a la información que sí tiene una estructura.
- La información no estructurada y su uso es importante para este sector de la economía y esto se comprueba con los datos obtenidos. Sin embargo, por ahora la gestión de esta información no es de gran relevancia para este negocio debido a que no se ha determinado cuál es la necesidad que ésta actividad tiene que satisfacer. Además, las empresas sin gestionar esta información siguen generando ingresos de tal manera que no afecta la viabilidad del negocio. Por otra parte, las empresas obtienen la mayoría de la información que agrega valor y apoya la toma de decisiones por medio de las bases de datos.
- El costo de estructurar el 100% de la información corporativa es inconmensurable en términos del hecho de pensar el modelo de datos y el esquema en el que se va a basar la respectiva información. Es por esto que la información no estructurada se gestiona de una manera diferente y es un problema complejo. Cabe resaltar que el costo no se da tanto en términos monetarios sino de capital y esfuerzo humano. Así mismo, es costoso porque se debe contar con una persona que tenga el conocimiento y la experiencia suficiente acerca de cada dominio

específico al que se refiera la información para poder generar un modelo de calidad a partir de su capacidad lógica y de análisis. Es posible añadir cierta estructura a partir de metadatos a los contenidos pero no se tiene el esquema riguroso que provee toda la semántica de la información.

- El sector del desarrollo de software está tratando de apoyar toda la toma de decisiones en datos, sin importar si la fuente es estructurada o no estructurada, dejando de lado los deseos, gustos o intuiciones personales.
- Algunas empresas están más adelantadas en la gestión de la información no estructurada que otras, debido a que tienen en producción un sistema de gestión documental que les ayuda a realizar ésta actividad.

Las siguientes conclusiones se refieren a la metodología utilizada en el trabajo de campo del estudio.

- Cuando el objeto de estudio es una organización se dificulta conseguir el espacio para realizar una entrevista personalizada.
- Como se trata de información que muchas empresas consideran confidencial, las personas entrevistadas se cohiben al momento de responder la guía de la encuesta.
- Debido a la dificultad de llevar el tema de estudio a términos cuantitativos, a medida que se consolidan datos se observó la necesidad de incluir mas preguntas.
- Actualmente existe muy poca bibliografía en libros acerca del tema de la información estructurada y no estructurada. Los aportes más importantes acerca de estos temas para la investigación se basaron en artículos de revistas especializadas de ciencias de la computación e

información de instituciones como la IEEE o IBM Research disponibles en línea o por medio de bases de datos en demostración.

13. RECOMENDACIONES

- Debido a la falta de conocimiento por parte de las empresas analizadas en este estudio, acerca de la información no estructurada y su procesamiento, se hace necesario plantear que los estudiantes de Ingeniería de Sistemas puedan acceder por medio de sus asignaturas obligatorias o de libre configuración, a una materia que exponga los conceptos acerca de la información no estructurada y su respectiva aplicación en la gestión de procesos y las herramientas informáticas. Esto debido a que el Ingeniero de Sistemas va a encontrar en el mercado laboral abundante información con poca o ninguna estructura que puede ser útil para la toma de decisiones. Con un conocimiento general del tema, esta persona estaría en total disposición de generarle a una organización, proyectos y aplicaciones para su beneficio y se incrementaría el valor agregado de los ingenieros de la universidad.
- Se recomienda crear un curso a través de Centros de Educación Continua para aquellas personas que estén relacionadas con el área y que no hayan tenido la oportunidad de asimilar los conceptos acerca de la información no estructurada y sus aplicaciones durante sus estudios formales. Esto con el fin de promover la importancia que tiene la información no estructurada y los beneficios de gestionarla adecuadamente.
- Indicarle a las empresas, a través de capacitaciones u otro canal de comunicación, la gran cantidad de información no estructurada que están manejando y las oportunidades empresariales que pueden generarse a partir de ésta, así como y la forma en que puede ser aprovechada.

- En el campo de la investigación en la Universidad se recomienda promover y continuar con estudios relacionados con la información no estructurada, por medio de tesis ó desarrollo de proyectos, ya se trata de un tema actual de investigación en las ciencias informáticas y que además puede generar un valor agregado a las empresas en los siguientes campos:

- Apoyo a la toma de decisiones.

- Mejora de los procesos.

- Competitividad.

A partir de este trabajo se generaron unas bases que muestran cual es la situación actual de las empresas Pyme Desarrolladoras de Software de la ciudad de Medellín. El alcance de las nuevas investigaciones deben abarcar este sector y otros sectores de la economía (civil, confecciones, inmobiliarias, etc.), enfocando también las grandes empresas y microempresas, para determinar las necesidades específicas de cada sector y poder proponer y desarrollar aplicaciones informáticas específicas.

BIBLIOGRAFÍA

[AN035] WIKIPEDIA. Animación [Documento electrónico]. Página Web. Estados Unidos: Artículo, 2007. (Citada: 18 noviembre 2007) <<http://es.wikipedia.org/wiki/Animaci%C3%B3n>>

[AP032] ABIAD, Pablo. Por primera vez admiten en el país al correo electrónico como prueba judicial. [Documento Electrónico]. Argentina: Diario el Clarín, miércoles 31 de octubre de 2001. (Citada: 18 Nov. 2007) <<http://www.redetel.gov.ar/Ambito%20Academico/Papers/novedad.htm>>

[BA005] BAEZA-YATES, Ricardo, EIRO-NETO, Berthier. "Modern Information Retrieval", Editorial Addison-Wesley, 1999.

[BC015] Gestión documental en las empresas. Bustelo Ruesta, Carlota. VII jornadas Españolas de Documentación, 2000.

[CA001] CASTAÑO, Adoración de Miguel., PIATTINI, Mario. "Fundamentos y modelos de bases de datos", Editorial Alfaomega, 1999.

[CC030] Cámara de Comercio de Medellín para Antioquia: Cotización Base de Datos. Información según la Cotización numero 86203 al 10 de Octubre de 2007.

[CO021] CARRIÓN, Juan. Conocimiento. [Documento Electrónico]: Página Web. (Citada 15 septiembre 2007) <http://www.gestiondelconocimiento.com/conceptos_conocimiento.htm>

[CO025] CONGRESO DE LA REPÚBLICA DE COLOMBIA. Ley 905 de 2004. [Documento Electrónico]: Página Web. Bogotá D.C.: Dada el 2 de Agosto de 2004 (Citada 10 agosto 2007) <http://www.sena.edu.co/NR/rdonlyres/8DCDB7D4-FABE-42DA-BD74-00F7C40AD9D0/0/Ley905_2004_MIPYMES.pdf>

[DA026] DANIES LACOUTURE, Rodolfo. Las mipymes en Colombia años 2002 – 2004. Superintendencia de Sociedades, Grupo de estadística Bogotá, 2005.

[DO029] IBM RESEARCH. Unstructured Information – The Knowledge Rush. [Documento Electrónico]: Página Web. (Citada 22 agosto 2007) <http://domino.research.ibm.com/comm/research_projects.nsf/pages/uima.knowledgeRush.html>

[FE006] FERRUCCI, D., LALLY A. “Building an example application with the Unstructured Information Management Architecture” IBM Systems Journal, Vol. 43, No 3, 2004.

[GE027] GENOVEVA RODRÍGUEZ, Astrid. La realidad de la Pyme Colombiana. [Documento Electrónico]: Página Web. Fundación Fundes, 2003 (Citada 15 agosto 2007) <<http://cec.uniandes.edu.co/docs/pyme.pdf>>

[GI011] GIL PECHUAN, Ignacio. “Sistemas y Tecnologías de la Información para la Gestión”. Mc Graw Hill. 1997

[GI033] GILLIS, Tom. Mejores prácticas para evitar la pérdida de información. [Documento electrónico]. Argentina: Auditoría Informática, 2007. (Citada: 18 noviembre 2007) <<http://www.riesgosinformaticos.com.ar/2007/11/04/mejores-practicas-para-evitar-la-perdida-de-informacion/>>

[IA020] WIKIPEDIA. Image file formats. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_file_formats>

WIKIPEDIA. Image analysis. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Computer_image_analysis>

WIKIPEDIA. Image search. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_search>

WIKIPEDIA. Image retrieval. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_retrieval>

WIKIPEDIA. Content-based image retrieval. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <<http://en.wikipedia.org/wiki/CBIR>>

[IN027] INFISA. "Buscamos desatar el poder de la información". [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <<http://www.infisaca.com/gestiondocumental/descripcion/descripcion.htm#ancla>>

[IT034] ITPYMES. El correo electrónico masivo hace caer la productividad laboral. [Documento electrónico]. Página Web. España: Artículo, 2007. (Citada: 18 noviembre 2007) <<http://itpymes.techtarget.es/desarrollo-personal/noticias/1000169014001/correo-electronico-masivo-fuente.1.html>>

[KE012] KEPNER, Charles H., TREGOE, Benjamin B. "El Directivo Racional". McGraw Hill. 1977.

[KN038] KNIGHT, Michael. The Continued Emergence of Enterprise Content Management. Febrero 2007. [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <http://www.contentmanager.net/magazine/article_1340_emergence_enterprise_content_management.html>

[LO036] LONDOÑO LONDOÑO, Luis Fernando. Recomendaciones para la Formación de una Empresa de Desarrollo de Software Competitiva en un País como Colombia. Centro de Investigación AVANSOFT S.A., Junio 2005.

[MA004] MALUF, David., BELL, David., ASHISH, Naveen., KNIGHT, Chris., TRAN, Peter. "Semi-structured Data Management in the Enterprise: A Nimble, High-Throughput, and Scalable Approach". Proceedings of the 9th International Database Engineering & Application Symposium, 2005.

[MA010] MARCH, James G. "Decisions and Organizations". Editorial Basil Blackwell. 1988.

[MA022] GESTIÓN ESTRATÉGICA DEL CONOCIMIENTO. MARTÍNEZ GARCÍA, Francisco Javier., PEÑALVER MARTÍNEZ, Antonio., SALAMANCA GARCÍA, Joaquín. [Documento Electrónico] (Citada: 15 sept. 2007) <<http://www.aui.org/archivos/gestionconocimiento.pdf>>

[MA040] MASON, Richard O., MITROFF, Ian I. A Program for Research on Management Information Systems. Management Science, Vol. 19, No. 5, Theory Series (Jan., 1973), pp. 475-487

[MI016] MICROSOFT ESPAÑA. Proteja a su empresa, proteja a sus empleados 2007 [documento electrónico]. (Citada: 10 Sep. 2007) <http://www.microsoft.com/spain/empresas/temas/correo_empresa/pautasinteligentes.msp>

[MK012] STONEBREAKER, Michael. Readings in Database Systems (Mogan Kaufmann Series in Data Management Systems)., Editorial: MORGAN KAUFMANN PUBLISHING; 3 edition, 1998

[MK019] MUNDO CONTACT S.A.. Polycom unifica el manejo de Video Conferencia y Contenido de Video. [Documento electrónico]. Página Web. Entrevista a Mukul Krishna, gerente global de la Práctica de Medios Digitales de Frost & Sullivan, 2007. (Citada: 18 noviembre 2007) <http://www.mundocontact.com/soluciones_detalle.php?recordID=1739>

[ML026] LAMARCA LAPUENTE, María Jesús. Hipertexto: El nuevo concepto de documento en la cultura de la imagen. Universidad Complutense de Madrid, Julio 2007 [Documento Electrónico]: Página Web. (Citada 20 agosto 2007) <http://www.hipertexto.info/documentos/b_datos.htm>

[MU023] MUDDU, Sudhakar. Unstructured Information. Septiembre 2005. [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <<http://www.s-ox.com/feature/article.cfm?articleID=1069>>

- [OC011] OBESO, Carlos. Información no estructurada. Agosto 2005. [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <http://www.wikilearning.com/informacion_no_estructurada-wkccp-3285-2.htm>
- [SE007] SENKO, M. E., ALTMAN, E. B., ASTRAHAN, M. M., FEHDER, P. L. "Data Structures and Accesing in Data-Base Systems I Evolution of Information Systems" IBM Systems Journal, No 1, 1973.
- [SE008] SENKO, M. E., ALTMAN, E. B., ASTRAHAN, M. M., FEHDER, P. L. "Data Structures and Accesing in Data-Base Systems II Information Organization" IBM Systems Journal, No 1, 1973.
- [ST024] STENSMO, Magnus., THORSON, Mikael. Unstructured Information Management. Publicado por Infosphere AB. Marzo 2003.
- [SU028] SARAWAGI, Sunita. Integrating unstructured data into relational databases. Proceedings of the 22nd international conference on data engeneering. 2006.
- [TW037] TWIKI. The Open Source Wiki for the Enterprise. [Documento Electrónico]: Página Web. Creado por Peter Thoeny. Última revisión noviembre 11 de 2007. (Citada 18 noviembre 2007) <<http://twiki.org/>>
- [UN018] MEROÑO CERDÁN, Angel Luis. El correo electrónico en las Pymes para la comunicación y gestión del conocimiento. [Documento electrónico]. Murcia: Universia España, 2005. (Citada: 18 noviembre 2007) <<http://www.universia.es/ubr/pdfs/UBR0012005070.pdf>>
- [VI014] HART, Jeffrey A. Power in the Information Age. Universidad de Indiana [Documento Electrónico]: Página Web. Bloomington: Marzo 1997 (Citada 10 agosto 2007) <<http://www.vii.org/papers/ciprut2.htm>>
- [WA003] WANG, K., HUIQING, L. "Discovering Structural Association Of Semistructured Data". IEEE Transactions On Knowledge And Data Engineering, Vol. 12, No. 3, May/June 2000.

[WE039] WETHERBE, James C. Executive Information Requirements: Getting It Right. MIS Quarterly, Vol. 15, No. 1 (Mar., 1991), pp. 51-65.

[WH002] WHITTEN, Jeffrey L., BENTLEY, Lonnie D., BARLOW, Victor M. Análisis y diseño de sistemas de información. Tercera Edición. McGraw Hill. 2000.

[WI031] WIKIPEDIA. Wiki. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 18 nov. 2007) <<http://es.wikipedia.org/wiki/Wiki>>

[WO009] C. WOOD, E. B. FERNANDEZ, R. C. SUMMERS. "Data Base Security: Requirements, policies and models" IBM Systems Journal, Vol. 19, No 2, 1980.

ALCANCE

Este estudio pretende diseñar y aplicar una encuesta sobre el uso y gestión de la información no estructurada en las empresas de desarrollo de software PYME de la ciudad de Medellín.

ESTUDIO SOBRE EL USO DE LA INFORMACIÓN NO ESTRUCTURADA EN LAS PYMES DE INGENIERIA DE SOFTWARE DE LA CIUDAD DE MEDELLÍN

Borja S. Juan Felipe, Guzmán Z. Julián, Lalinde P. Juan Guillermo, *Universidad EAFIT*

Abstract—Este artículo proporciona información acerca del uso de la información no estructurada en las pymes de la ciudad de Medellín. Se hace un recorrido por los temas principales que hacen parte del estudio como también un análisis referente a los puntos más importantes que se generaron a partir del trabajo de campo. Al final se proporcionan las conclusiones que se extrajeron de este estudio.

Términos clave—Información estructurada, información no estructurada, gestión de la información, toma de decisiones.

I. INTRODUCCIÓN

El proceso de toma de decisiones en una empresa se ve influenciado por información, en la cual se basan las personas para seleccionar entre varias alternativas, la más conveniente. Las bases de datos relacionales tienen mucha importancia en este punto, ya que de estas se puede proveer de información de primera categoría a las personas. Sin embargo, la información tiene tres tipos según el contexto del presente estudio, a saber: estructurada, semiestructurada y no estructurada.

Por ahora, resulta adecuado afirmar que la información que reside en las bases de datos relacionales es estructurada y que el resto pertenece a las otras dos categorías. Es decir, en contraste la información no estructurada la podemos encontrar en documentos de texto como correos u hojas de cálculo y también en objetos tipo mapa de bits como las imágenes, video y audio. Esta información es proporcionalmente, mas abundante que la estructurada y su tasa de crecimiento es aun mayor. [DO029] y [KN038]

En este trabajo se observará, en primera instancia, si las empresas de desarrollo de software de la ciudad de Medellín están utilizando la información no estructurada y cuál es el tratamiento que estas empresas le están dando a la misma. Estos datos provienen de la mejor técnica para obtener datos cualitativos y cuantitativos: entrevistas acompañadas de una encuesta guía que se le realiza a las diferentes empresas de la ciudad de Medellín. Se pretende así, recoger datos acerca de los diferentes tipos de información no estructurada que se manejan a nivel organizacional (audio, video, imágenes y texto

incluyendo el correo electrónico), las áreas, cargos involucrados y bajo qué métodos. Además se podrá evaluar cual es el nivel de conocimiento que las empresas tienen de la información no estructurada a través del uso de herramientas para la gestión de ésta y datos porcentuales involucrando las bases de datos y tomas de decisiones que se hacen con respecto a este tipo de información.

II. OBJETIVOS

OBJETIVO GENERAL

Analizar cómo gestionan y cómo usan la información no estructurada en la toma de decisiones las Pymes de la ciudad de Medellín en el sector de desarrollo de software.

OBJETIVOS ESPECÍFICOS

1. Determinar el tamaño de las muestras necesarias para que el estudio sea significativo.
2. Establecer la proporción de uso de la información no estructurada en las empresas del sector objeto de estudio.
3. Determinar las formas de uso más comunes de la información no estructurada en las empresas del sector objeto de estudio.
4. Identificar las principales herramientas que se utilizan actualmente para el uso de la información no estructurada en las empresas del sector objeto de estudio.
5. Determinar el tipo de información no estructurada existente en las pymes del sector objeto de estudio en Medellín
6. Determinar si las Pymes del sector objeto de estudio en Medellín gestionan la información no estructurada.

7. Determinar la influencia de la Información no estructurada en la toma de decisiones de las Pymes del sector objeto de estudio en Medellín.

III. MARCO TEÓRICO

INFORMACIÓN ESTRUCTURADA

Es la información que ha sido analizada y organizada. Tiene un esquema que describe los datos y que proviene de un modelo de datos. Por lo general se almacena en una base de datos, motor de búsqueda o base de conocimiento. Como afirma David Ferrucci y Adam Lally “La información estructurada está caracterizada como la información cuyo significado intencional no es ambiguo y está explícitamente representado en la estructura o formato de los datos. El ejemplo canónico de la información estructurada es la tabla de la base de datos relacional”. [FE006] Cada elemento de la información está asociado con un esquema preciso y bien definido donde cada encabezado de una columna de la base de datos especifica su semántica, definiendo cómo la información debería ser interpretada por una aplicación informática o por un usuario final.

Se había mencionado anteriormente que la información está compuesta por hechos y al relacionarlo con las representaciones, M. Senko afirma que “La información consiste de hechos acerca de las cosas. Estos hechos y cosas existen independientemente de cualquier representación, pero es esencialmente imposible operar conceptualmente con ellos excepto en términos de alguna representación”. [SE008] Acerca de estas representaciones, M. E. Senko especifica que “las representaciones son materializaciones concretas de la información estructurada. [...] Los computadores trabajan directamente con las representaciones y sólo indirectamente, si mucho, con la información representadas por ellas.” [SE008]

INFORMACIÓN SEMIESTRUCTURADA

La información semiestructurada es aquella en donde los datos no se rigen por un esquema estricto de una base de datos, es decir, su estructura no es regular. Son datos que pueden ser irregulares o estar incompletos y tienen una estructura que puede cambiar rápida o impredeciblemente. La información semiestructurada está constituida en esencia, por documentos, tales como documentos de texto y presentaciones en diferentes formatos. “Tal información es típicamente semiestructurada, esto es, que existe alguna estructura en los documentos pero no exactamente una estructura formal impuesta por un esquema de bases de datos”. [MA004]

Según Ke Wang y Huiqing Liu: “La información semiestructurada se presenta cuando la fuente de los datos no impone una estructura rígida (tal como la Web) y cuando los datos provienen de fuentes heterogéneas. A diferencia de la información no estructurada, la semiestructurada si tiene algún tipo de estructura.” [WA003]

INFORMACIÓN NO ESTRUCTURADA

Son diferentes las definiciones en este campo, pero cuando se refiere a información no estructurada, esencialmente se trata de todos los contenidos almacenados de los cuales no se puede decir específicamente como están organizados, como están relacionados con el resto de la información o como deben ser usados. En esencia, hacen parte de esta categoría los objetos tipo mapa de bits, los cuales no están basados en ningún lenguaje natural (imágenes, video y audio) y los archivos de texto, que están basados en lenguaje escrito, tales como el correo electrónico. En esta categoría se clasifican todos los documentos para los cuales no hay meta datos semánticos.

La información estructurada puede ser manejada con soluciones que soportan “*queries*” y reportes a tipos predeterminados de datos y relaciones previstas. La información no estructurada no tiene una definición conceptual o de tipo de datos.

INFORMACION NO ESTRUCTURADA EN LAS EMPRESAS

Documentos

En las empresas se pueden observar dos clases de documentos: Electrónicos que son todos los que están en una archivo digital (Word, Excel, pdf, etc.) y tangibles que son todos aquellos documentos que están en papel (notas, balances generales, hojas de vida, etc.). Los archivos en papel todavía tienen una gran implementación en las organizaciones. Mucha parte de la información sigue siendo archivada de manera tangible produciendo graves problemas a la hora de su gestión y por el espacio que estos ocupan.

A pesar de que cada día se están generando más documentos electrónicos, Carlota Bustelo Ruesta nos cuenta: “la acumulación de documentos electrónicos en los ordenadores de las empresas aparece como un nuevo problema que hasta ahora no existía. El espacio en este caso no preocupa demasiado, pues la memoria de almacenamiento es cada vez más barata; pero la dificultad para encontrar lo que se busca y la generación de copias de seguridad empiezan a ser considerados problemas importantes de gestión.” [BC015]

Claro que esto no sucede solamente para los documentos electrónicos, en el caso de los archivos físicos también se generan grandes problemas a la hora de querer conseguir un documento específico.

Sin embargo, se hace indispensable que para ambos casos exista una gestión documental que sirva de apoyo al personal y al negocio como tal. Los objetivos básicos se pueden definir

de la siguiente manera teniendo en cuenta la necesidad de las empresas:

- “Facilitar a las personas trabajar con los documentos. Cada persona sabe qué documentos tiene que guardar, cuando, como y donde. Cada persona sabe como encontrar en poco tiempo los documentos adecuados cuando los necesita.
- Facilitar que la información se comparta y se aproveche como un recurso colectivo, evitar que se duplique, evitar fotocopios innecesarios, evitar dobles grabaciones de datos, etc.
- Conservar la memoria de la organización más allá de los individuos que trabajan en ella y poder aprovechar el valor de los contenidos en los que queda plasmada la experiencia, evitando empezar de cero sobre aspectos en los que ya hay experiencia acumulada.” [BC015]

Correos Electrónicos

Microsoft, en su página web española, destaca: “Sin lugar a dudas, Internet se ha convertido en la herramienta de comunicaciones empresariales más utilizada. Según algunos cálculos, su inmensa popularidad ha generado 275 millones de buzones corporativos. Otras estadísticas indican que el 60 por ciento de la información empresarial más importante se transmite en la actualidad por correo electrónico”. [MI016]

Es tal el impacto que el correo electrónico ha generado en las empresas que inclusive éste se puede volver bastante peligroso. Esto debido a que mucha información importante esta viajando por este medio y puede ser robada en cualquier momento. Además, se ha llegado a ver, que los mismos correos electrónicos pueden ser utilizados como pruebas en las cortes judiciales.

Según una encuesta realizada por Harris Interactive (2003): “El email es la tecnología favorita seguida del teléfono móvil, los portátiles, el teléfono, el procesador de textos, la hoja de cálculo, el fax y otros.” [UN018] Es indispensable comentar, según la misma encuesta, que el 83,7% de las empresas tienen como herramienta el correo electrónico.

Esta tecnología ha sido muy acogida por los empresas ya que involucra un bajo costo, tiene una amplia difusión (clientes, proveedores, público, empresas, etc.), facilidad de uso, diversidad de usos (envío de archivos, agenda, búsqueda, etc.). [UN018] Además de la comunicación que se puede obtener con cualquier persona a nivel mundial.

Por todo esto, se hace clave llevar un control de los emails implantando políticas para el uso del correo electrónico en el trabajo, así como una serie de principios que regulen el uso de esta tecnología.

Audio, Video y Animaciones:

“El contenido de video está proliferando en las empresas e incluye video conferencias grabadas, transmisiones por la Web, capacitación de los empleados, direcciones de los ejecutivos, y más. Las empresas se están esforzando para lidiar con este volumen de contenido cada vez mayor mientras mejoran el video como un factor estratégico. Los productos de puntos han forzado tradicionalmente a los clientes a unir las soluciones separadas”. [MK019]

El video en las organizaciones puede ser utilizado de diferentes formas. Entre ellas las videoconferencias que se realizan para tomar decisiones con personas ubicadas en diferentes lugares geográficos, así mismo, puede utilizarse para concretar negocios y validar propuestas. Entre otras también podemos encontrar: Relaciones con los inversionistas, anuncios internos, reuniones de todos los empleados, fusiones y adquisiciones, etc. En la mayoría de casos, estas videoconferencias no suelen quedar archivadas, se realizan en vivo sin generar ningún almacenamiento para su posterior utilización. Por otro lado, el video ayuda en las capacitaciones que se le dan a los empleados al ingresar a las empresas, e inclusive, para generar conocimiento de una herramienta dada. Esto se puede evidenciar mediante retención de la propiedad intelectual, aprendizaje a distancia, conferencias grabadas, programas de estudio individual e inclusive certificaciones para los empleados. El contenido de video también puede ayudar a reducir los costos de soporte del cliente al proporcionar detalles de los productos, instrucciones de instalaciones, guías para ubicación y resolución de problemas y manuales de usuarios. Cabe destacar que en este mismo campo, también están las animaciones. [AN035] Estas animaciones cumplen un papel similar para las empresas. Por lo general, las animaciones se crean en FLASH y sirven de soporte para las capacitaciones, presentaciones de sus productos, campañas de mercadeo, promociones, etc. ya sea en la página Web o para vender sus productos directamente.

Imágenes

Son diversas las extensiones de imágenes que se pueden manejar a nivel corporativo. Esto depende en gran medida a la unidad de negocios que cada empresa maneja dentro de su portafolio de productos o servicios. Es diferente la organización, mantenimiento y gestión que le pueda dar una organización de diarios nacionales a la que pueda darle una empresa promotora en salud. Sin embargo, para ambos es de vital importancia la información que allí se pueda generar.

A parte de todos las extensiones que se manejan (.jpg, gif, .bmp, .png, .tiff, etc.) existen otras, no electrónicas, que se elevan en el nivel de importancia para las empresas. Una imagen puede ser: una radiografía, un caso de uso, un logo en papel, un mapa e inclusive una señal de tránsito. Todas ellas cumpliendo un papel fundamental para cada organización.

Cabe resaltar que ha medida que transcurre el tiempo se hace de mayor importancia tener todas las imágenes en medio

electrónico, en gran medida, a que las imágenes tangibles pueden sufrir de instalaciones inadecuadas que condicionan su conservación, demoras en las búsquedas (por ser manual) y problemas para su traslado.

Para el caso de las empresas de desarrollo de software, gran parte del negocio está sustentado en la elaboración de diagramas. Para esto se utilizan diversas herramientas que ayudan a su elaboración pero no a su organización. No podemos dejar de lado todas aquellas imágenes que aún se desarrollan y archivan en papel.

Entre otras utilizaciones puntuales de las imágenes podemos encontrar: el logo de la empresa, imágenes para capacitaciones, explicaciones de productos, mercadotecnia, etc. [IA020]

Conocimiento

En las empresas existe otra fuente de información no estructurada, la cual reside en el conocimiento de las personas. Al respecto, Juan Carrión afirma que:

"El Conocimiento es un conjunto integrado por información, reglas, interpretaciones y conexiones puestas dentro de un contexto y de una experiencia, que ha sucedido dentro de una organización, bien de una forma general o personal. El conocimiento sólo puede residir dentro de un conocedor, una persona determinada que lo interioriza racional o irracionalmente". [CO021]

Este mismo autor nos afirma que este conocimiento influye en el diario accionar de las personas, tomando decisiones acerca de los fenómenos percibidos, además, que es uno de los intangibles que permite obtener una ventaja competitiva. Para este último punto, Francisco Martínez aclara que "el conocimiento se transforma en un aprendizaje continuo que es vital para el funcionamiento sostenible de las organizaciones" [MA022] relacionado con las competencias, con la capacidad de innovar y mejorar, y con el compromiso y la motivación.

GESTIÓN DE LA INFORMACIÓN NO ESTRUCTURADA

Según Sudhakar Muddu, "la información no estructurada es información que por lo general no ha sido gestionada. Este tipo de información se almacena de diversas formas y se crean individual y manualmente más que automáticamente. El sistema de archivos en el cual ésta información reside no se supervisa típicamente y el contenido es prácticamente invisible a los interesados. La búsqueda es generalmente una búsqueda indexada." [MU023]

Según Magnus Stensmo y Mikael Thorson, "La Gestión de la Información no Estructurada consiste en las herramientas y métodos necesarios para almacenar, acceder y recuperar, navegar y generar conocimiento primordialmente de la información basada en texto", definición que se adopta para el presente estudio. [ST024]

IV. ANÁLISIS DE DATOS

¿Utilizan videos o animaciones dentro o fuera de la empresa para explicar procesos o procedimientos?

El 62% (13 empresas) entrevistado respondió que si utilizan videos o animaciones para explicar algún proceso o procedimiento perteneciente a la compañía. De esta parte el 46% (6 empresas) afirma que se utiliza con el fin de capacitar a sus empleados en el manejo de las herramientas y métodos comunes en el proceso de desarrollo de software. El 38.46% afirma que se utiliza para la presentación de sus productos a los clientes. Finalmente, en una menor cantidad, se utiliza con el fin de promocionar la empresa.

El 38% (8 empresas) restante respondió que no utiliza ninguna animación ni video.

¿Se utilizan imágenes o diagramas para soportar algún proceso?

El 76% (16 empresas) entrevistadas respondió que si utilizan imágenes o diagramas en algún proceso o procedimiento perteneciente a la compañía. Todas las empresas utilizan un método de modelación en su proceso de desarrollo de software. Pero también se genera información no estructurada a partir de las siguientes variables: sistema de gestión y calidad de la empresa, imágenes de las aplicaciones para capacitación de los clientes, cronograma de actividades y gestión de proyectos, imágenes de las aplicaciones para comunicaciones internas, modelación y documentación de procesos que ayudan a las diferentes certificaciones, como también la posibilidad de digitalizar documentos físicos debido a su importancia para el negocio como: facturas, hojas de vida, visitas, contratos tanto de clientes como de empleados, quejas y reclamos, correspondencia entrante y saliente.

El 24% (5 empresas) restante respondió que no utilizan imágenes en sus procesos.

Pero solo un 25% (4 empresas) tiene herramientas apropiadas para la gestión de las imágenes como un sistema de gestión documental. Esto implica que hay un 75% (12 empresas) que simplemente tienen la información (imágenes en este caso) almacenadas sin tener algún tipo de método que les ayude a gestionar este tipo de información.

¿Se utiliza audio para soportar algún proceso?

El 48% (10 empresas) entrevistado respondió que si utilizan audio en algún proceso o procedimiento perteneciente a la compañía, en especial aquellos procesos donde hay un contacto con los clientes y en los cuales se quiere dejar una constancia de lo que se expresó y se arregló entre las partes. Es decir, esto tiene que ver con la definición del dominio del producto a desarrollar, los contratos y el levantamiento de requisitos. También se utiliza el audio como parte de otro documento para explicar algún proceso o dar una capacitación. El 52% (11 empresas) restante respondió que no utiliza audio.

¿Qué herramientas informáticas utiliza para el apoyo de los procesos en la empresa?

Dentro de las principales herramientas que generan información, ya sea estructurada o no estructurada, encontramos las siguientes: entornos de interfaces de desarrollo, herramientas de modelado de software, sistemas operativos, CRM's (Customer Relationship Management), motores de bases de datos, contabilidad, facturación, gestión de proyectos, exploradores de Internet, aplicaciones para copias de seguridad, herramientas desarrolladas para su manejo interno como aplicaciones para el manejo de hojas de vidas.

Es importante resaltar que solo el 19% (4 empresas) utilizan un sistema de gestión documental ya que esta herramienta permite manejar cierta información no estructurada, en su mayoría textos, de manera indexada.

¿De esas herramientas cuáles generan información que agreguen valor a la empresa?

A pesar de que el proceso central de la población objeto de estudio es el de desarrollo de software, el 43% (9 empresas) considera que todas sus herramientas informáticas generan información que agrega valor a la empresa. Además, afirman que en caso de que se cumpla lo contrario serían eliminadas. Un 5% (1 empresa) aclara que las herramientas informáticas no son las que generan información que agregue valor sino que son las personas las encargadas de esta función luego de agregarle criterio a los datos arrojados por las aplicaciones.

Un 62% (13 empresas), incluyendo las empresas que respondieron que todas las herramientas generan valor, indicaron que las aplicaciones de desarrollo también lo hacen. Así mismo, un 19% (4 empresas), sin incluir las que respondieron que todas, advierte que las aplicaciones que generan indicadores, consolidaciones y reportes son las que permiten agregar valor porque facilitan el seguimiento de los procesos. Por último, el 24% (5 empresas) de las empresas respondieron que las aplicaciones de gestión de proyectos son las más importantes.

¿Tiene forma de observar las relaciones que hay entre el correo electrónico, las imágenes, el portal Web, texto, video, audio, papeles tangibles?

Un 33% (7 empresas) de los entrevistados respondieron que si pueden observar relaciones entre los contenidos. Sin embargo, estas relaciones no se pueden obtener, básicamente, a través de una aplicación sino que se realiza por medio de una operación manual. Para ello se utilizan matrices de trazabilidad, proceso que se toma bastante tiempo y que puede resultar tedioso, tanto para los empleados que lo realizan como para la obtención inmediata de resultados para la organización. También se pueden clasificar todos los contenidos en categorías que dependen de un evento que los relaciona, que a su vez, va de la mano del almacenamiento en estructura de directorios.

Por el contrario, el 57% (12 empresas) afirma que no es posible relacionar estos contenidos. Inclusive, las respuestas se tornan un tanto más próximas a la relación que hay entre los directorios y sus subdirectorios y no acerca del contenido de esos directorios en sí.

Mientras que un 10% (2 Empresas) no sabe o no responde si tiene forma de observar las relaciones entre los contenidos.

¿Qué conoce acerca de la Información no estructurada?

El 10% (2 Empresas) entrevistadas, además de citar características y propiedades del tema, estuvieron acertadas en dar una definición teórica del mismo y en estos casos tenían desarrollada una herramienta para tal efecto.

El 19% (4 Empresas) conocen poco acerca de este tema. Algunas de las definiciones recibidas tuvieron que ver con características de la información no estructurada pero no con su definición formal. Entre las respuestas más comunes se encuentran las siguientes:

- Es una información que tiene difícil ubicación e indexación.
- Es información que se puede encontrar en papel, documentos office, conversaciones telefónicas, faxes, etc.
- Algoritmos de búsquedas de cosas no previstas.
- Comentarios de clientes como cosas que se dicen y que se oyen, cosas de la prensa que no se registran en ningún medio.

Los resultados generales muestran que el 71% (15 Empresas) de las organizaciones no conocen nada acerca de la información no estructurada. De hecho, el término generó en las personas entrevistadas respuestas inapropiadas que inclusive no se asemejan, de ninguna forma, a la respuesta esperada. A su vez, generó reacciones negativas y de mucha incertidumbre para continuar con la entrevista. Y esta es una de las razones por la que la pregunta se encuentra al final de la guía de la encuesta. Inclusive, muchas personas devolvieron la pregunta buscando tener una idea con respecto a este tema.

¿Qué porcentaje de información que maneja es no estructurada?

A diferencia de la pregunta relacionada con la información estructurada en donde se podía observar a simple vista una tendencia mayoritaria cercana al 100%, en esta, se observa que no existe ninguna tendencia hacia un valor en particular. De hecho, no existen muchas diferencias entre los intervalos presentados.

Cabe resaltar, que las personas encuestadas al no conocer acerca de la información no estructurada, respondió en un 24% (5 empresas) que no saben cuanto porcentaje estimado manejan de este tipo de información.

V. CONCLUSIONES

- Las empresas que tengan un conocimiento medio apropiado de la información no estructurada tienen un mayor campo de acción a la hora de tomar decisiones que generen valor para la empresa. Esto sin incluir que las empresas, teóricamente hablando, no tienen un gran conocimiento acerca de este tipo de información. Sin embargo, son capaces de llevarlo a la práctica y evidenciar gran ventaja con respecto a sus competidores.
- Todos los procesos empresariales se vuelven mucho más ágiles, se disminuye el tiempo de ejecución, búsqueda y almacenamiento en cada uno de los departamentos pertenecientes a la compañía.
- Se evidenció que el correo electrónico es una de las fuentes que más genera información no estructurada en las empresas de este sector, de la misma forma los documentos también generan un alto índice de este tipo de información. En un segundo plano aparecen las imágenes seguidas por el video y el audio. A pesar de que todas las empresas reciben llamadas telefónicas, realizan contactos con los clientes para definición de requerimientos, reuniones para el seguimiento del producto y contratos, no se respalda esta información en audio. El portal interno ocupa un lugar importante ya que está continuamente en crecimiento y modificación por parte de los empleados con nuevas versiones que incluyen más contenidos. Por último está el portal Web donde debido a sus pocas actualizaciones no genera mucha información no estructurada.
- Los usos más comunes que se le dan al correo electrónico son las comunicaciones con los clientes, entre los mismos empleados o con los proveedores. El audio para el levantamiento de requisitos y la definición de contratos. El video para la capacitación de empleados en algún método o herramienta, para la capacitación de los clientes con un producto nuevo y para la presentación de la empresa. Las imágenes se usan para el proceso de ingeniería, el modelado de procesos en el sistema de calidad y también para capacitaciones y comunicaciones internas. El portal interno se utiliza para documentar los procesos de la empresa, los proyectos, las aplicaciones, herramientas y métodos.
- La parte de la información no estructurada sobre la que se tiene menos control es el correo electrónico. Este fenómeno se da ya que las empresas manipulan el correo como sucesos aislados y además, no hacen parte de ningún sistema de gestión documental.
- La gran mayoría de las empresas no utilizan una herramienta que les permita gestionar la información no estructurada de una manera adecuada. Hasta ahora, pocas empresas recurren a un Sistema de Gestión Documental que facilite esta actividad de una manera más automática. Estas herramientas están enfocadas al manejo de archivos basados en lenguaje natural.
- En las empresas que gestionan la información no estructurada se puede ver un beneficio en la toma de decisiones tan pertinente como si la fuente de la información proviniese de un esquema estructurado.
- En la pequeña y mediana industria del desarrollo del software de la ciudad de Medellín se tiene la sensación de tener toda la información bajo una estructura de la cual se puede obtener todo su significado. Una de las razones por las que se puede explicar este fenómeno es por el desconocimiento de qué es la información no estructurada. Por otro lado, porque la información que se encuentra en las bases de datos relacionales es la de vital importancia para el funcionamiento de este negocio. Por último, debido a que la parte documental de las organizaciones también hace parte de las bases de datos pero no se tiene clara la diferencia con respecto a la información que sí tiene una estructura.
- La información no estructurada y su uso es importante para este sector de la economía y esto se comprueba con los datos obtenidos. Sin embargo, por ahora la gestión de esta información no es de gran relevancia para este negocio debido a que no se ha determinado cuál es la necesidad que ésta actividad tiene que satisfacer. Además, las empresas sin gestionar esta información siguen generando ingresos de tal manera que no afecta la viabilidad del negocio. Por otra parte, las empresas obtienen la mayoría de la información que agrega valor y apoya la toma de decisiones por medio de las bases de datos.
- El costo de estructurar el 100% de la información corporativa es inconmensurable en términos del hecho de pensar el modelo de datos y el esquema en el que se va a basar la respectiva información. Es por esto que la información no estructurada se gestiona de una manera diferente y es un problema complejo. Cabe resaltar que el costo no se da tanto en términos monetarios sino de capital y esfuerzo humano. Es posible añadir cierta estructura a partir de meta datos a los contenidos pero no se tiene el esquema riguroso que provee toda la semántica de la información.
- El sector del desarrollo de software está tratando de apoyar toda la toma de decisiones en datos, sin importar si la fuente es estructurada o no

estructurada, dejando de lado los deseos, gustos o intuiciones personales.

Las siguientes conclusiones se refieren a la metodología utilizada en el trabajo de campo del estudio.

- Cuando el objeto de estudio es una organización se dificulta conseguir el espacio para realizar una entrevista personalizada.
- Como se trata de información que muchas empresas consideran confidencial, las personas entrevistadas se cohiben al momento de responder la guía de la encuesta.

Debido a la dificultad de llevar el tema de estudio a términos cuantitativos, a medida que se consolidan datos se observó la necesidad de incluir mas preguntas.

VI. RECOMENDACIONES

- Debido a la falta de conocimiento por parte de las empresas analizadas en este estudio, acerca de la información no estructurada y su procesamiento, se hace necesario plantear que los estudiantes de Ingeniería de Sistemas puedan acceder por medio de sus asignaturas obligatorias o de libre configuración, a una materia que exponga los conceptos acerca de la información no estructurada y su respectiva aplicación en la gestión de procesos y las herramientas informáticas. Esto debido a que el Ingeniero de Sistemas va a encontrar en el mercado laboral abundante información con poca o ninguna estructura que puede ser útil para la toma de decisiones. Con un conocimiento general del tema, esta persona estaría en total disposición de generarle a una organización, proyectos y aplicaciones para su beneficio y se incrementaría el valor agregado de los ingenieros de la universidad.
- Se recomienda crear un curso a través de Centros de Educación Continua para aquellas personas que estén relacionadas con el área y que no hayan tenido la oportunidad de asimilar los conceptos acerca de la información no estructurada y sus aplicaciones durante sus estudios formales. Esto con el fin de promover la importancia que tiene la información no estructurada y los beneficios de gestionarla adecuadamente.
- Indicarle a las empresas, a través de capacitaciones u otro canal de comunicación, la gran cantidad de información no estructurada que están manejando y las oportunidades empresariales que pueden generarse a partir de ésta, así como y la forma en que puede ser aprovechada.
- En el campo de la investigación en la Universidad se

recomienda promover y continuar con estudios relacionados con la información no estructurada, por medio de tesis ó desarrollo de proyectos, ya se trata de un tema actual de investigación en las ciencias informáticas y que además puede generar un valor agregado a las empresas en los siguientes campos:

- Apoyo a la toma de decisiones.
- Mejora de los procesos.
- Competitividad.

A partir de este trabajo se generaron unas bases que muestran cual es la situación actual de las empresas Pyme Desarrolladoras de Software de la ciudad de Medellín. El alcance de las nuevas investigaciones deben abarcar este sector y otros sectores de la economía (civil, confecciones, inmobiliarias, etc.), enfocando también las grandes empresas y microempresas, para determinar las necesidades específicas de cada sector y poder proponer y desarrollar aplicaciones informáticas específicas.

BIBLIOGRAFIA

[AN035] WIKIPEDIA. Animación [Documento electrónico]. Página Web. Estados Unidos: Artículo, 2007. (Citada: 18 noviembre 2007) <<http://es.wikipedia.org/wiki/Animaci%C3%B3n>>

[BC015] Gestión documental en las empresas. Bustelo Ruesta, Carlota. VII jornadas Españolas de Documentación, 2000.

[CO021] CARRIÓN, Juan. Conocimiento. [Documento Electrónico]: Página Web. (Citada 15 septiembre 2007) <http://www.gestiondelconocimiento.com/conceptos_conocimiento.htm>

[DO029] IBM RESEARCH. Unstructured Information – The Knowledge Rush. [Documento Electrónico]: Página Web. (Citada 22 agosto 2007) <http://domino.research.ibm.com/comm/research_projects.nsf/pages/uima.knowledgeRush.html>

[FE006] FERRUCCI, D., LALLY A. “Building an example application with the Unstructured Information Management Architecture” IBM Systems Journal, Vol. 43, No 3, 2004.

[IA020] WIKIPEDIA. Image file formats. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_file_formats>

WIKIPEDIA. Image analysis. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007)

<http://en.wikipedia.org/wiki/Computer_image_analysis>

WIKIPEDIA. Image search. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_search>

WIKIPEDIA. Image retrieval. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <http://en.wikipedia.org/wiki/Image_retrieval>

WIKIPEDIA. Content-based image retrieval. [Documento Electrónico]: Página Web. Última Actualización noviembre de 2007 (Citada 20 agosto 2007) <<http://en.wikipedia.org/wiki/CBIR>>

[KN038] KNIGHT, Michael. The Continued Emergence of Enterprise Content Management. Febrero 2007. [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <http://www.contentmanager.net/magazine/article_1340_emergence_enterprise_content_management.html>

[MA004] MALUF, David., BELL, David., ASHISH, Naveen., KNIGHT, Chris., TRAN, Peter. "Semi-structured Data Management in the Enterprise: A Nimble, High-Throughput, and Scalable Approach". Proceedings of the 9th International Database Engineering & Application Symposium, 2005.

[MA022] GESTIÓN ESTRATÉGICA DEL CONOCIMIENTO. MARTÍNEZ GARCÍA, Francisco Javier., PEÑALVER MARTÍNEZ, Antonio., SALAMANCA GARCÍA, Joaquín. [Documento Electrónico] (Citada: 15 sept. 2007) <<http://www.aui.org/archivos/gestionconocimiento.pdf>>

[MI016] MICROSOFT ESPAÑA. Proteja a su empresa, proteja a sus empleados 2007 [documento electrónico]. (Citada: 10 Sep. 2007) <http://www.microsoft.com/spain/empresas/temas/correo_empresa/pautasinteligentes.msp>

[MK019] MUNDO CONTACT S.A.. Polycom unifica el manejo de Video Conferencia y Contenido de Video. [Documento electrónico]. Página Web. Entrevista a Mukul Krishna, gerente global de la Práctica de Medios Digitales de Frost & Sullivan, 2007. (Citada: 18 noviembre 2007) <http://www.mundocontact.com/soluciones_detalle.php?recordID=1739>

[MU023] MUDDU, Sudhakar. Unstructured Information. Septiembre 2005. [Documento Electrónico]: Página Web. (Citada 10 agosto 2007) <<http://www.sox.com/feature/article.cfm?articleID=1069>>

[SE008] SENKO, M. E., ALTMAN, E. B., ASTRAHAN, M. M., FEHDER, P. L. "Data Structures and Accessing in Data-Base Systems II Information Organization" IBM Systems Journal, No 1, 1973.

[ST024] STENSMO, Magnus., THORSON, Mikael. Unstructured Information Management. Publicado por Infosphere AB. Marzo 2003.

[UN018] MEROÑO CERDÁN, Angel Luis. El correo electrónico en las Pymes para la comunicación y gestión del conocimiento. [Documento electrónico]. Murcia: Universia España, 2005. (Citada: 18 noviembre 2007) <<http://www.universia.es/ubr/pdfs/UBR0012005070.pdf>>

[WA003] WANG, K., HUIQING, L. "Discovering Structural Association Of Semistructured Data". IEEE Transactions On Knowledge And Data Engineering, Vol. 12, No. 3, May/June 2000.