

Almacenes de datos aplicados a la seguridad ciudadana

Public Safety Data Warehouse

Asnioby Hernández López¹, Iván M. Cárdenas Tandrón², Yanet Peña Vazquez¹

¹Universidad de Ciencias Informáticas (UCI), {ahernandezlo, ypenav}@uci.cu

²Universidad Central Marta Abreu de Las Villas (UCLV), ivanmct@uclv.edu.cu

Resumen

En los países de Latinoamérica, uno de los principales problemas que poseen los Gobiernos es que se encuentra limitada su capacidad informativa y analítica para proyectar estrategias integrales de prevención e intervención en materia de seguridad ciudadana. Es necesario establecer un modelo iterativo e incremental para constituir las bases del desarrollo del Almacén de Seguridad Ciudadana, enfocado en la asimilación paulatina de la información por parte de los analistas, así como en el incremento del dominio de análisis del propio almacén. Esta información, al ser presentada de forma dinámica al usuario final, le permite a este, poder combinarla con distintos parámetros y alcanzar mejores interpretaciones útiles a la institución en apoyo a la toma de decisiones.

Palabras clave: almacenes de datos, integración, inteligencia institucional, seguridad ciudadana.

Abstract

In the countries of Latin America, one of the main problems is that governments have limited its ability to project information and analytical strategies, comprehensive prevention and intervention in matters of public safety. It is necessary to establish a model for iterative and incremental constitute the foundation for the development of Public Safety Warehouse, focused on the gradual assimilation of information by analysts as well as an increase in the domain analysis own warehouse. This information, to be presented in a dynamic way to the end user, allows this, can be combined with different parameters and achieve better interpretations useful to the institution in supporting decision-making.

Key words: business intelligence, citizen security, data integration, data warehouse.

Introducción

Debe entenderse como Seguridad Ciudadana el grado de respeto que se otorga al conjunto de derechos de los ciudadanos, no solo por parte del Estado, sino también de parte de las personas e instituciones públicas y privadas, que deben garantizar el bienestar de todos los componentes de las estructuras a las cuales están vinculadas, a las cuales prestan servicios o de las que depende su seguridad [Ríos, 2006].

De eso se desprende que toda persona tiene derecho a la protección por parte del Estado a través de los órganos de seguridad ciudadana regulados por ley, frente a situaciones que constituyan amenaza, vulnerabilidad o riesgo para la integridad física de las personas, sus propiedades, el disfrute de sus derechos y el cumplimiento de sus deberes.

En los países de Latinoamérica, uno de los principales problemas que poseen los Gobiernos es que se encuentra limitada su capacidad informativa y analítica para proyectar estrategias integrales de prevención e intervención en materia de seguridad ciudadana.

Luego, se hace necesario relacionar las informaciones contenidas en las distintas fuentes de información empleando herramientas informáticas que propicien su integración, transformación y almacenamiento, de tal manera que el análisis de esta información integrada apoye la toma de decisiones en la formulación de medidas, así como en el seguimiento y control de éstas desde una perspectiva proactiva y de corresponsabilidad ciudadana.

Una solución altamente eficiente y validada son los Almacenes de Datos que convierten los datos operacionales de una organización en una herramienta profesional, haciéndolos disponibles a los usuarios que lo necesiten para el análisis y toma de decisiones [Thomsen, 2002]. Según Bill H. Inmon, padre reconocido del *Data Warehousing*, “El Data Warehouse (DWH) es una colección de datos orientados al tema, integrados, no volátiles e historizados que facilitan la toma de decisiones” [Inmon, 2005].

Metodología de diseño

La metodología de diseño utilizada se basa en la definición de un Almacén de Datos aplicado a temas de Seguridad Ciudadana, donde se persigue resolver los requerimientos de información que existen en los Ministerios Públicos, que hoy día tienen limitada su capacidad informativa y analítica para proyectar estrategias integrales de prevención e intervención en materia de seguridad ciudadana.

El desarrollo de un Almacén de Datos implica identificar y conceptualizar el objeto de análisis, las fuentes legadas que contienen la información requerida, las estructuras físicas que tendrá el almacén, así como las reglas de extracción, transformación y carga de los datos a estas estructuras físicas.

A continuación se presenta la evolución histórica de los sistemas gestores de bases de datos, hasta llegar a los almacenes de datos. Y la justificación de por qué su uso en el Ministerio Público.

Avances tecnológicos de los sistemas gestores de bases de datos

El modelo relacional tiene entre sus objetivos guardar la integridad de los datos obtenidos en los procesos transaccionales automatizados (OLTP: Procesamiento Transaccional en Línea). Sin embargo, este modelo no se corresponde con la forma en la que el usuario percibe la gestión del conocimiento de una institución, en general, y la gestión pública, en particular. De hecho, Codd afirmó que aunque los sistemas de gestión de bases de datos relacionales han sido muy beneficiosos para los usuarios, nunca han sido diseñados para proporcionar funciones potentes de síntesis, análisis y consolidación de los datos [Codd, 1990].

A partir de mediados de los ochenta, en el entorno empresarial, ha cobrado importancia el concepto *Data Warehouse* o Almacén de Datos (AD), entendido como una colección de datos con determinadas cualidades que se estructuran

para apoyar el proceso de toma de decisiones.

Existen numerosas definiciones de *Data Warehouse*, si bien la más conocida fue propuesta por Inmon. [Inmon, 1996] “Un AD es una colección de datos orientados a temas, integrados, no-volátiles y variante en el tiempo, organizados para soportar necesidades empresariales”.

Hay que señalar que el diseño del AD no es un proceso trivial, se debe elegir, en base a la información que se desea explotar, los datos que se guardarán, la unidad mínima de éstos, la estructura de las entidades de información, las dimensiones que se estudiarán, estadísticos intermedios que se deben conservar y muchos aspectos más para que el diseño responda a las necesidades de información de distintos departamentos o áreas y niveles jerárquicos de la Institución, así como la eficiencia en la provisión operacional de dicha información [Manjit, 2000].

En este sentido, las bases de datos que conforman el componente de almacenamiento físico del AD se caracterizan por los siguientes aspectos:

1 Integradas: deben constituir un conjunto de datos y metadatos perfectamente integrados con respecto al nombre de las variables, formatos de los distintos campos, medida de los atributos, codificación, etc.

2 Temáticas: las bases de datos deben conformarse hacia materias o temas, como clientes, productos, campañas, etc., a diferencia de las bases de datos de los sistemas operacionales, más orientados a procesos administrativos.

3 Históricas: éste es un factor clave en la toma de decisiones, contar con información histórica para comparar datos en distintos períodos e identificar tendencias. El tiempo debe estar en todos y cada uno de los registros del DW, de manera que, cuando un dato entra en el DW se sepa en qué momento tenía ese valor.

4 No Volátiles: la información una vez incorporada al DW debe mantenerse, en general, invariable, cargándose una vez en el tiempo y no permitiendo actualizaciones de los datos.

Justificación del almacén de datos en la gestión pública

Es válido destacar que los Almacenes de Datos no deben ser considerados una simple copia de los datos generados por la Organización o entes externos a esta. A continuación se exponen algunas razones que justifican su adopción como una plataforma viable ante la creciente demanda de información de las personas tomadoras de decisiones.

Algunos de los aspectos que deben someterse a consideración se comentan a continuación:

1 Rendimiento: el tiempo de acceso y consulta de datos se reduce drásticamente, debido a que todos los datos se encuentran en un repositorio y no es necesario el acceso a plataformas heterogéneas para su obtención. Los datos se almacenan en estructuras optimizadas para responder a consultas complejas y no previstas.

2 Combinación de Orígenes de Datos: contar con datos de diferentes fuentes internas o externas aporta la visión necesaria para tomar mejores decisiones pero esta tarea, en la mayoría de los casos, resulta compleja y requiere tiempo a las personas encargadas de tomar decisiones. Por ejemplo, es probable que no se utilicen los mismos

criterios de almacenamiento (nombres de las entidades, atributos considerados, tipos etc.) en las bases de datos de distintos departamentos. Sin embargo, en el Almacén de Datos los datos se homogenizan durante el proceso de carga y disponen de la calidad requerida para su análisis posterior.

3Limpieza de los datos: En las Instituciones coexisten múltiples Sistemas de Información; aumentando su complejidad la posibilidad de utilizar varios de ellos para atender un mismo conjunto de operaciones, y en esos casos es probable que los datos se dupliquen, contengan errores, estén incompletos, etc. El AD corrige estos inconvenientes en el proceso de carga para evitar interpretaciones erróneas de los mismos.

4Disponibilidad de datos históricos: Los datos históricos suelen desecharse porque no son de utilidad en los Sistemas de Gestión que tienen como objetivo representar el estado actual de cada Área de la Institución, pero son un elemento esencial de cualquier análisis. El Almacén de Datos es el lugar adecuado para estos datos.

5Datos Resumidos y Agregados: Muchas veces para tomar decisiones, no es necesario incluir los datos detallados, por lo que en el AD solo estarán disponibles con el nivel de detalle requerido (por ejemplo en temas de Seguridad Ciudadana los homicidios por estados, los operativos policiales mas efectivos, indicadores de victimización, etc.).

El Almacén de Datos constituye una herramienta capaz de almacenar, estructurar y hacer disponibles los datos útiles en el proceso de toma de de decisiones.

Resultados y Discusión

El modelo propuesto para el desarrollo de un Almacén de Datos dicta algunos pasos y consideraciones a tener en cuenta para su desarrollo y el caso de los Almacenes para la Seguridad Ciudadana no es un caso atípico en este sentido. Además, se presenta una breve discusión de las principales diferencias existentes entre la Inteligencia de Negocios y la Inteligencia Institucional.

Integración de Fuentes de información

Para el proceso de integración de las fuentes es necesario establecer un modelo iterativo e incremental de 4 pasos que constituye las bases del desarrollo del Almacén de Seguridad Ciudadana, enfocado en la asimilación paulatina de la información por parte de los analistas, así como en el incremento del dominio de análisis del propio almacén. Dicho modelo prevé la mejora continua de la organización en cuanto a la calidad de la gestión.

1	Se requiere de Nueva Información para ampliar el dominio de análisis.
2	Identificación y clasificación de Fuentes Fiables de obtención de la Información.
3	Proceso de Integración.
4	Análisis y explotación sistemática, por parte de los analistas, del nuevo dominio informativo ampliado que puede conllevar a la necesidad de adquisición de nueva información.

Vale destacar que antes de proceder con los Procesos de Integración es necesario conceptualizar los términos y valores de las variables a recibir de las fuentes de información.

A diferencia de Almacenes de datos aplicados a entornos empresariales, las fuentes de información tienen un alto grado de autonomía que conduce a que algunos términos y el conjunto de valores de las mismas variables sean diferentes. Aunque se supone que las Instituciones del Gobierno, Órganos de Seguridad y otras, deben seguir una única línea en cuanto a las informaciones comunes que manejan, en muchos casos no sucede así.

Es común ver como la Policía y las instituciones de Salud, por ejemplo, hacen alusión a los mismos elementos, de manera diferente.

Esto es uno de los principales problemas en la Integración de Entidades que brindan información realmente fiable para resolver problemas como el que tratamos. En entornos empresariales este fenómeno aparece en menor medida, puesto que todos los departamentos, o Sub Entidades de una Corporación, llevando al caso extremo, concuerdan en manejar los conceptos del negocio de forma análoga, en la mayoría de los casos.

Inteligencia institucional

El término Inteligencia Organizacional o Institucional existe ya desde hace algún tiempo y se utilizó por primera vez en un informe de Gartner Group en el año 1996 como Inteligencia de Negocios, pero sus bases han sido conocidas mucho antes de esa fecha. En aquel tiempo, estos principios fueron conocidos como Gestión de Sistemas de Información y Sistemas de Soporte de Negocios. Sin importar el nombre, los principios son los mismos.

En la actualidad, según el *Data Warehousing Institute*, en el ámbito empresarial, al conjunto de procesos, tecnologías y herramientas que se necesitan para convertir los datos en información útil para el apoyo a la toma de decisiones se le denomina Inteligencia de Negocios.

Si se parte de que el funcionamiento de un Sistema de Información (SI) contempla cuatro actividades fundamentales, puede decirse que la primera es la entrada de datos, los que son almacenados, procesados y constituyen salidas en forma de información, que al convertirse en conocimiento del recurso humano, y explotado en planes que buscan un retorno de inversión, generan un nuevo cúmulo de datos, con lo que forman nuevamente una entrada al SI, y se completa el ciclo.

Es en la actividad de procesamiento de datos donde interfieren los procesos, tecnologías y herramientas para el análisis y presentación de los datos. Esta información, al ser presentada de forma dinámica al usuario final, le permite a este, poder combinarla con distintos parámetros y alcanzar mejores interpretaciones útiles a la empresa en apoyo a la toma de decisiones, con lo que se mejora el retorno de inversión de la empresa, se logra un conocimiento de sus clientes, de sus procesos internos y su organización.

Inteligencia de Negocios vs Inteligencia Institucional

Los indicadores para medir el desempeño de las organizaciones en cada caso no son iguales. La diferencia más

importante entre la Inteligencia de Negocios y la Inteligencia Institucional es la medición del Retorno de Inversión (ROI por sus siglas en inglés). Para las instituciones gubernamentales, el ROI no es únicamente un balance de resultados financieros.

Su medición tiene indicadores tales como: más y mejor educación, más y mejores empleos, o reducción de la pobreza, por citar algunos ejemplos. De ahí que los datos que sirven para medir el desempeño de los gobiernos no se encuentran en las oficinas públicas, sino en la calle, en las ciudades y pueblos, o de dicho de otro modo, en la satisfacción ciudadana y en una mejor calidad de vida para todos.

Una segunda gran diferencia entre la implementación de una estrategia de Inteligencia de Negocios y una Inteligencia Institucional, es la heterogeneidad de las fuentes de información. Para una empresa, cruzar diferentes datos sobre su operación puede resultar complejo, pero cuando se trata de organismos gubernamentales las fuentes son mucho más diversas. El gobierno es la suma de una gran variedad de dependencias, organismos, empresas paraestatales, organizaciones sociales y ciudadanas, entre otras, que trabajan en distinto ámbitos territoriales [Rivera, 2007].

La Inteligencia Institucional es un concepto que se utiliza mayormente en las Instituciones de Gobierno, que para enfrentar situaciones sociales, utilizan las mismas técnicas que la Inteligencia de Negocios, con la diferencia de que su resultado no se mide en un retorno de inversión monetario, sino en soluciones sociales, como pueden ser disminución de los delitos, mejoras en la calidad de vida del ciudadano o sencillamente en la percepción de seguridad que tienen las personas [Rivera, 2007].

Conclusiones

Los almacenes de datos aplicados a la seguridad ciudadana facilitan la integración, y explotación de la información de los diferentes Órganos de Seguridad Ciudadana. Asimismo, permiten el análisis y presentación de estudios multidisciplinarios de forma dinámica, lo que conlleva a la rápida obtención de resultados a través de los cuales puedan identificarse patrones o tendencias de comportamiento en situaciones determinadas.

Referencias Bibliográficas

- [Armstrong-Smith, 2006] Darlene and Michel Armstrong-Smith. Oracle Discoverer 10g Handbook. San Francisco, California, The McGraw-Hill Companies, Inc, 2006.
- [Clark, Waterhouse, Fletcher, 2004] Mike Clark, Mark Waterhouse, Piter Fletcher, Web Services Business Strategies and Architectures, USA, Ed Wiley, Inc, 2004.
- [Codd, 1990] Codd, E. F. (1990). - The relational Model for Database Management: version 2, Addison-Wesley Publishing Company, Reading.
- [Imber, 2000] Imber, Jane. Dictionary of Marketing Terms, 3rd Edition, Hauppauge, NY Barron's, 2000.
- [Inmon, 1996] Inmon, William H. (1996). - Building the Data Warehouse, Wiley Computer, New York.
- [Inmon, 2005] W. H. Inmon (2005). Building Data Warehouse, Fourth Edition. USA. John Wiley & Sons, Inc.

- [Kemper, 2007] Nicolas Kemper Valverde. Minería de Datos como Herramienta para la Gestión Moderna, 2007
- [Keogh, 2003] Jim Keogh, J2EE Manual de Referencia, Interamericana de México, SAU. The McGraw-Hill Companies, Inc, 2003.
- [Manjit, 2000] Manjit, S. (2000).- Developing a Corporate Data Warehousing Strategy Enterprise System Integration, Chapter 33, CRC Press LLC, Boca Ratón, Florida, 449-467.
- [Martínez, 2002] Valle, Marco A. (2002)- Coordinador Nacional Proyecto de Seguridad Ciudadana, Programa de las Naciones Unidas para el Desarrollo (PNUD), Nicaragua. [Disponible en: <http://www.iigov.org/ss/article.drt>]
- [O'Brien, Marakas, 2006] James O'Brien and George Marakas, Management Information Systems, 7th ed. The McGraw-Hill Companies, Inc, 2006.
- [Ríos, 2006] Ríos, Wilfredo. Trabajo de Diploma presentado para optar por el título de Ingeniero en Ciencias Informáticas. Año 2006. Universidad de Ciencias Informáticas.
- [Thomsen, 2002] Thomsen, Erick (2002) OLAP Solutions, Building Multidimensional Information Systems, Second Edition. Jhon Wiley & Sons, Inc. USA.
- [Rivera, 2007] Rivera Victoria, Sofía Mónica. El Datawarehouse y el Business Intelligence en Gobierno. 2007. [Disponible en: <http://www.monografias.com>]