

“Fortalecimiento del Programa Nacional de Monitoreo
Atmosférico: Extensión de la cobertura del Sistema Nacional
de Calidad del Aire (SINAICA)”

Reporte final

JICA LIBRARY



1177862181

Matías Software Group S.A. de C.V.

Noviembre de 2004

JICA
615
61.9
MXO
LIBRARY

MXO
JR
04-02

Índice

Antecedentes.....	3
Consideraciones.....	3
Actividades.....	5
Productos.....	7
1.Documentación sobre las características del tipo y formato de los archivos de datos o de la tecnología de almacenamiento utilizada en todas las redes de monitoreo atmosférico automáticas.....	7
2.Programa de instalación de la modificación a los módulos extractores.....	9
3.Código fuente completo de las modificaciones y nuevos desarrollos, relación de programas desarrollados y bibliotecas utilizadas durante el desarrollo de los mismos. En todos los casos el proveedor deberá otorgar al INE/CENICA la licencia no exclusiva que permita tanto el uso como de distribución ilimitada, incluso de carácter comercial, de todos los programas desarrollados.....	9
4.Reporte conteniendo la documentación de los procedimientos de instalación y operación de los programas de la Tercera Versión del sistema. Este reporte deberá incluir: Scripts de creación y mantenimiento de la base de datos según se requiera, parámetros de compilación, en su caso, de los programas y documentación de configuración de los mismos, incluyendo secuencia de instalación, diagramas del flujo de datos y de procedimientos, diccionario de datos y diagrama entidad/relación de la base de datos. Además de los nuevos algoritmos de los programas de graficación utilizados en el sistema para conocer el criterio de interpretación de los datos.....	10
5.Instalación de un mínimo de 14 sistemas colectores y transmisores de datos de los centros de control de las redes al SINAICA.....	10
6.Módulos para el manejo de datos de: a) las redes manuales de monitoreo atmosférico, b) Equipo de monitoreo y análisis de contaminantes atmosféricos del CENICA: equipo de monitoreo continuo de COVs, estaciones de monitoreo automático, analizador de carbón orgánico y globos piloto; para su posible integración al SINAICA. Ver anexo I [de los términos de referencia] para descripción detallada de los módulos para equipos del CENICA.....	11
7.Módulo extractor SQL tanto para Windows como para Linux.....	12
8.Módulo extractor de comunicación directa con aquellos dataloggers que cuenten con documentación de los protocolos de comunicación utilizados.....	12
9.Módulos de consulta y reportes de datos en tiempo real y de consulta especializada. Este último deberá incorporar tecnologías de Sistemas de Información Geográfica (SIG) para mostrar la información en forma de mapas, por ejemplo donde se muestren isolíneas de concentración, mapas de ozono, etc.....	12
10.Extensiones al subsistema extractor o de almacenamiento para la validación automática de los datos de los centros de control antes de aparecer en el subsistema de presentación.....	12
11.Sistema de generación y administración de certificados digitales X509 como “Autoridad Certificadora Central” sin limitaciones en el tipo y cantidad de certificados generables.....	13
12.15 copias del reporte final en español y 15 copias del reporte final en inglés	13
Comentarios y recomendaciones.....	13



1177862【8】

Extensión de la cobertura del SINAICA

Anexo 1.....	15
Introducción.....	15
El sistema Colector/Transmisor.....	17
El formato DBF.....	20
El formato de E-DAS de ESC.....	27
El formato de la RDBMS en SQL.....	34
Anexo 2.....	36
Introducción.....	36
Sistema Colector y Transmisor de Dato autónomo.....	37
Administración remota.....	38
Anexo 3.....	39
Introducción.....	39
Plataformas.....	39
Requerimientos.....	39
Parámetros de compilación.....	40
Relación de programas desarrollados y bibliotecas utilizadas.....	43
Anexo 4.....	46
Licencia.....	46
Anexo 5.....	56
Introducción.....	56
Compilación y ordenamiento de las bases de datos existentes.....	56
Interfase teodolito-base de datos.....	57
Generación de vectores de viento.....	57
Aplicación de análisis y reporte.....	58

Reporte Final

Antecedentes.

A finales del año 2002 Matías Software Group S.A. de C.V. desarrolló un prototipo demostrativo que permitió automatizar el proceso de recepción de datos de la calidad del aire de las redes de monitoreo automáticas de las ciudades de México, Guadalajara y Toluca mediante el uso de los protocolos TCP/IP dadas las actuales facilidades en el acceso a Internet.

Si bien el prototipo del Sistema Nacional de Información de la Calidad del Aire (SINAICA) resultó funcional en su momento, esa versión requirió de ampliaciones y de un proceso de modificaciones acordes con las pautas de modernización actuales para la difusión de información y de los cambios en los equipos y sistemas de adquisición de datos de las redes de monitoreo atmosférico.

Así, hacia el segundo semestre del 2003 se plantearon mejoras al prototipo señalado y se desarrolló la segunda versión del sistema de adquisición de datos del SINAICA, con el que se resolvieron problemas detectados en el envío de datos de las redes hacia el servidor de base de datos y se hicieron modificaciones a la página de presentación de resultados.

Es importante mencionar que en esta segunda versión se incorporó a la Red Estatal de Monitoreo Atmosférico de la ciudad de Puebla dentro del SINAICA, por lo que en ese momento se tuvieron cuatro redes conectadas al sistema: Guadalajara, México, Puebla y Toluca; además de que durante el desarrollo de la segunda versión del sistema se hicieron diagnósticos preliminares de la infraestructura informática y de los procesos de información para la incorporación de las redes de Monterrey y Salamanca.

Consideraciones

Partiendo de la segunda versión del SINAICA desarrollada en 2003, se identificaron algunos elementos necesarios para la extensión de la cobertura del SINAICA, así como posibles mejoras al subsistema de consulta y presentación de datos.

Extensión de la cobertura del SINAICA

Estos elementos quedaron plasmados en los términos de referencia del contrato que se transcriben a continuación:

“... La incorporación de dos redes: Monterrey y Salamanca. Dado que la primera ha pasado por un proceso de modernización tanto de los instrumentos de medición, como de los sistemas de captura y respaldo de información, no se cuenta ahora con la decodificación de los datos de la misma, la cual se encuentra planteada en una estructura de programación mediante el lenguaje SQL, por lo que es necesaria de forma adicional al programa de comunicación, un módulo extractor específico. Con respecto a la Red de Salamanca, si bien no cuenta con una estructura de datos compleja, dado que recientemente se han tomado muestras de ellas, es necesario la elaboración del módulo de comunicación para la misma, esta red concentrará su información en enero de 2004 en las oficinas de gobierno del Instituto de Ecología de Guanajuato, de donde se enlazará al SINAICA.

Se ha identificado que las redes de Irapuato, Celaya, Cd. Juárez, Tijuana, Mexicali, Aguascalientes, Torreón, Zacatecas y San Luis Potosí, tienen amplias posibilidades de ser integradas al SINAICA.”

“Así mismo, se plantea integrar al SINAICA la red de monitoreo automático del CENICA, con sus dos estaciones automáticas de referencia.

Identificar las condiciones necesarias para integrar al SINAICA la información generada por las redes de monitoreo manuales y plantear las soluciones.”

“Por otra parte, es un hecho que el programa de reportes tal como ha sido formulado por el consultor funciona de forma efectiva. Sin embargo, de la segunda versión se han identificado una cantidad importante de renovaciones y mejoras al aspecto, mismo que se encuentra disponible en la página Web del Instituto Nacional de Ecología: <http://sinaica.ine.gob.mx>, y que ha raíz de esto se han recibido sugerencias de mejoras de los diferentes usuarios y de los responsables operativos de las redes. Por lo que ahora se plantea el desarrollo de nuevas herramientas de consulta para adecuarse a las necesidades de otros tipos de usuarios, lo que representa desarrollar diferentes módulos de presentación de datos y generación de reportes de acuerdo a los requerimientos de cada tipo de usuario:

Para el público en general la información debe ser presentada de forma

Extensión de la cobertura del SINAICA

amigable y en términos claros, en todo momento tener presente la comparación contra las normas de calidad del aire y lo que ellas significan. Se requiere que el módulo permita la consulta de dos o más estaciones y de más de dos parámetros de forma simultánea.

Usuarios especializados requieren tanto de las bases de datos como de un resumen de datos históricos con gráficos estadísticos y por otra parte, las condiciones actuales también con parámetros y gráficos estadísticos que permitan un rápido y mayor grado de asimilación de los mismos.

Tomadores de decisión o también llamados administradores de calidad de aire, tener indicadores de calidad del aire en forma de reporte, del tipo resumen ejecutivo.

Integrar en la base de datos del SINAICA los datos no solo crudos (sin validar) sino también los ya validados por las redes y ponerlos a disposición a los diferentes usuarios, entre ellos las diferentes áreas del INE y SEMARNAT así como gobiernos locales, grupos de investigación, etc. Se plantea que la nueva versión debe considerar el envío periódico automatizado de los datos validados por las redes y la conformación de la base de datos validados.”

“Un aspecto por demás relevante en la segunda versión, es la preocupación por parte de las redes a ser cuestionados sobre la veracidad de los datos, tal como está implementado en la actualidad, el SINAICA recibe datos crudos, sin validar, tan pronto cómo se colectan en las redes, es decir, no se cuenta con un filtro que las proteja, ya que la información recibida en un determinado momento puede no necesariamente reflejar condiciones ambientales reales, sino ser producto de la calibración de un instrumento, mal funcionamiento de los monitores, limitaciones de los sistemas de transferencia de datos propios de cada de la Red u otros factores. Las redes de Monitoreo de la Ciudad de México, Guadalajara y Puebla cuentan con aplicaciones o programas que les permiten utilizar este tipo de validación primaria, evitando la difusión de ese tipo de información a los usuarios. Por lo que en éste proyecto es importante incorporar al SINAICA los mecanismos de filtrado que hagan ésta validación primaria para la presentación de información al usuario final.”

Actividades

En los términos de referencia se identificaron nueve actividades que resultan en doce productos.

Extensión de la cobertura del SINAICA

Para efectos del presente informe las actividades se han correlacionado con los productos solicitados en virtud de que no siempre resulta evidente su relación.

En la tabla a continuación se listan las actividades tal como aparecen en los términos de referencia anotándose el número que identifica a los productos correspondientes. En los casos en que la actividad es de índole genérica o consideración de diseño (no relacionada con algún producto en particular) esto se indica con la abreviatura 'GEN'.

# Actv	Descripción	# Prod
1	Realizar un diagnóstico de las características de las redes de monitoreo de calidad del aire, manuales y automáticas para determinar los protocolos específicos necesarios para la extracción y transmisión de información al SINAICA	1
2	Integrar a las redes de monitoreo automáticas de Monterrey, Salamanca, Irapuato, Celaya, Cd. Juárez, Tijuana, Mexicali, CENICA mas aquellas que al término del contrato y que por su formato y equipamiento pudieran integrarse al sistema.	5
3	Diseñar y desarrollar un sistema colector y transmisor de datos que permita unificar los requerimientos de los Centros de Control de las Redes.	2,7,8
4	Desarrollar un módulo (Módulo de Administración) que permita el monitoreo del sistema colector y transmisor de datos mencionado en el punto anterior, en forma remota y desde las instalaciones del INE/CENICA.	2
5	El proveedor deberá asegurar el mejor acceso para la transmisión de datos durante el primer año pudiendo utilizar desde enlaces dedicados a Internet hasta sistema dialup e incluyendo la comunicación entre el INE y CENICA. Será responsabilidad de las redes dar las facilidades para la instalación de esta infraestructura.	GEN
6	El proveedor deberá incorporar la capacidad de que la transmisión de datos se realice de forma segura mediante el uso de certificados digitales X509, debiendo proporcionar al INE/CENICA todas las herramientas para la generación y administración de los mismos.	11
7	Capacitar al personal del INE/CENICA incluyéndose la instalación, operación y adecuación de las aplicaciones en los centros de control y estaciones y sistemas de almacenamiento del INE.	GEN

Extensión de la cobertura del SINAICA

# Actv	Descripción	# Prod
8	Elaborar aplicaciones para el manejo de datos de las redes manuales así como de los equipos de monitoreo y análisis de contaminantes atmosféricos del CENICA para la integración de la información generada por éstos al SINAICA. Las actividades detalladas respecto a las aplicaciones necesarias para los equipos de monitoreo y análisis del CENICA se incluyen en el Anexo 1 .	6
9	Modificar los subsistemas de almacenamiento y de presentación existentes para que incluya las siguientes características. preferentemente en forma modular: Consulta en línea en tiempo real (el tiempo más próximo) Consulta especializada Validación local Datos históricos. Validación federal	9

Productos

De las actividades anotadas en los términos de referencia se desprendieron diferentes productos, a continuación se transcriben en cursivas seguidos de los resultados obtenidos al final del proyecto, refiriéndose a los anexos de este documento cuando su extensión así lo requiere.

En virtud de que muchos de los productos se materializaron en desarrollo de software estos se han instalado según su naturaleza ya sea en el nuevo servidor de bases de datos, en equipos propios de las redes de monitoreo, equipamiento de CENICA o en el firmware del módulo extractor y colector de datos, adicionalmente se han concentrado en un disco compacto, mismo del que se ha entregado el original a la DGCENICA. En esos casos el hecho se reporta en el detalle del producto correspondiente.

1. Documentación sobre las características del tipo y formato de los archivos de datos o de la tecnología de almacenamiento utilizada en todas las redes de monitoreo atmosférico automáticas.

Con objeto de dar cumplimiento a la actividad general número uno y con objeto de determinar el tipo y formato de los archivos de datos o tecnología utilizados por las redes se realizaron varias actividades puntuales en las redes de monitoreo, que la tabla a continuación resume:

Extensión de la cobertura del SINAICA

<i>Red</i>	<i>Fecha de actividad</i>	<i>Actividad general</i>
CJU	13/abr 29/abr al 1/may 22/jul 23/jul al 25/jul	Conferencia telefónica Evaluación y Diagnóstico Conferencia telefónica Integración al SINAICA
GDL		Integración al SINAICA
MTY	31/mar 17/jun 26/jul	Evaluación y diagnóstico Prueba Piloto Integración al SINAICA
MXC	25/jun	Evaluación y Diagnóstico Integración al SINAICA
PUE	30/ago	Integración al SINAICA
ROS	26/jun	Evaluación y Diagnóstico Integración al SINAICA
BAJIO/GTO	14/abr	Visita Centro de control GTO
BAJIO/LEO	15/abr	Evaluación y Diagnóstico
BAJIO/IRA	15/abr	Evaluación y Diagnóstico
BAJIO/SAL	16/abr	Evaluación y Diagnóstico Integración al SINAICA
TIJ	23/jun 27/jun	Evaluación y Diagnóstico Pruebas Piloto Integración al SINAICA
TKT	26/jun	Evaluación y Diagnóstico Integración al SINAICA
TLC		Evaluación y Diagnóstico Integración al SINAICA

Extensión de la cobertura del SINAICA

En el **Anexo 1** se describen los resultados del diagnóstico practicado así como los tipos de archivos y formatos detectados con todo detalle.

2. Programa de instalación de la modificación a los módulos extractores.

Desde las etapas tempranas del proyecto y de conformidad con lo establecido en las consideraciones generales y en particular con la actividad número tres se determinó que la mejor estrategia era diseñar un sistema colector y transmisor de datos autónomo.

El **Anexo 2**, sobre la implementación del sistema, dedica un capítulo a este desarrollo.

3. Código fuente completo de las modificaciones y nuevos desarrollos, relación de programas desarrollados y bibliotecas utilizadas durante el desarrollo de los mismos. En todos los casos el proveedor deberá otorgar al INE/CENICA la licencia no exclusiva que permita tanto el uso como de distribución ilimitada, incluso de carácter comercial, de todos los programas desarrollados.

En el disco compacto se incluye todo el código fuente.

En el **Anexo 3** de Detalles Técnicos, se encuentra un capítulo con la relación de programas desarrollados y bibliotecas utilizadas.

Por la naturaleza del desarrollo realizado, la plataforma seleccionada y el espíritu de cooperación que ha guiado desde sus orígenes el desarrollo del SINAICA, Matías Software Group considera que la GPL (General Public License) de la FSF (Free Software Foundation) es la licencia que al mismo tiempo de satisfacer los requerimientos solicitados, garantiza y fomenta la participación activa de toda la comunidad interesada en el desarrollo futuro del SINAICA, por lo que otorga no sólo al INE/CENICA ésta licencia de forma irrevocable sino a cualquier parte interesada en el presente o en el futuro.

En el **Anexo 4** se incluye el texto completo de licencia tal como la presenta la FSF.

Extensión de la cobertura del SINAICA

- 4. Reporte conteniendo la documentación de los procedimientos de instalación y operación de los programas de la Tercera Versión del sistema. Este reporte deberá incluir: Scripts de creación y mantenimiento de la base de datos según se requiera, parámetros de compilación, en su caso, de los programas y documentación de configuración de los mismos, incluyendo secuencia de instalación, diagramas del flujo de datos y de procedimientos, diccionario de datos y diagrama entidad/relación de la base de datos. Además de los nuevos algoritmos de los programas de graficación utilizados en el sistema para conocer el criterio de interpretación de los datos.**

Ver Anexo 3.

- 5. Instalación de un mínimo de 14 sistemas colectores y transmisores de datos de los centros de control de las redes al SINAICA.**

Como se documenta en el **Anexo 1**, no todas las ciudades planteadas contaban con una red como tal, menos con un centro de control, así que se tuvieron que implementar estrategias alternas de conectividad, lo que permitió finalmente la integración a esta versión del SINAICA de las ciudades de Guadalajara, Ciudad Juárez, Mexicali, Ciudad de México, Monterrey, Puebla, Rosarito, Salamanca, Tecate, Tijuana, Toluca y la red experimental de CENICA.

En la tabla que aparece más arriba se muestran las fechas en que las diferentes ciudades se fueron incorporando a la nueva versión del SINAICA.

Es importante hacer notar que diversos problemas con la infraestructura de comunicaciones no resueltos en tiempo impidieron finalmente la incorporación de las ciudades de Celaya e Irapuato al sistema lo que motivó en su momento que se hicieran adecuaciones a los alcances del proyecto. Sin embargo el sistema está listo para su incorporación y CENICA cuenta con los sistemas colectores correspondientes.

Extensión de la cobertura del SINAICA

- 6. Módulos para el manejo de datos de: a) las redes manuales de monitoreo atmosférico, b) Equipo de monitoreo y análisis de contaminantes atmosféricos del CENICA: equipo de monitoreo continuo de COVs, estaciones de monitoreo automático, analizador de carbón orgánico y globos piloto; para su posible integración al SINAICA. Ver anexo 1 [de los términos de referencia] para descripción detallada de los módulos para equipos del CENICA.**

Las dos estaciones que forman la red de CENICA se incorporaron directamente al SINAICA con sus respectivos Colectores y Transmisores de Datos de forma que todos los mecanismos de consulta funcionan de la misma forma que para el resto de las redes, adicionalmente se cargaron al sistema los datos históricos disponibles.

Es importante señalar que los datos de la estación de Revolución no se encuentran aún disponibles al público en general, pues están en proceso de revisión por parte de CENICA.

En el disco compacto se incluyen tres de los cuatro módulos desarrollados.

Para el desarrollo correspondiente al monitoreo continuo de Compuestos Orgánicos Volátiles (COVs) se programó un conjunto de macros para Excel, ya instalados en el equipo correspondiente. Por ser estos de uso específico para los datos ya capturados no se incluyó en el disco compacto.

La aplicación para el analizador de carbón orgánico y elemental se desarrolló como un sistema independiente al SINAICA y consiste en una interfase hacia la base de datos nativa de dicho sistema (MS Access) y un módulo de post procesamiento y calibración.

El desarrollo correspondiente a los globos piloto tiene dos partes, una para automatizar la captura, que se conecta directamente al teodolito utilizado y una parte de consulta integrada al sistema general del consultas del SINAICA (en un módulo independiente), su documentación completa se encuentra en el **Anexo 5**.

Por último, el sistema para fluorescencia de rayos X se integró como otro módulo

Extensión de la cobertura del SINAICA

independiente al sistema general del consultas del SINAICA.

7. Módulo extractor SQL tanto para Windows como para Linux

Este módulo, requerido en la actualidad por la red de la ciudad de Monterrey, se incorporó al desarrollo del módulo colector y transmisor de datos autónomo detallado en el **Anexo 2** y se encuentra en operación en la red citada desde el momento de su incorporación al SINAICA, el 26 de julio.

8. Módulo extractor de comunicación directa con aquellos dataloggers que cuenten con documentación de los protocolos de comunicación utilizados.

Este módulo, desarrollado originalmente para los dataloggers de CENICA, se encuentra en operación en todas las estaciones de las ciudades del estado de Baja California incorporado al Sistema Colector y Transmisor de Datos autónomo.

9. Módulos de consulta y reportes de datos en tiempo real y de consulta especializada. Este último deberá incorporar tecnologías de Sistemas de Información Geográfica (SIG) para mostrar la información en forma de mapas, por ejemplo donde se muestren isolíneas de concentración, mapas de ozono, etc.

Todos los desarrollos relativos a consulta y reportes se encuentran instalados en el nuevo servidor de CENICA, accesible desde Internet en la dirección del Web <http://sinaica2.ine.gob.mx/> además de incluirse en el disco compacto.

10. Extensiones al subsistema extractor o de almacenamiento para la validación automática de los datos de los centros de control antes de aparecer en el subsistema de presentación.

Estos desarrollos se encuentran incorporados en parte al Sistema Colector y Transmisor de Datos autónomo y en parte a los sistemas de almacenamiento en el nuevo servidor de CENICA, además de incluidos en su totalidad en el disco compartio.

Extensión de la cobertura del SINAICA

11. Sistema de generación y administración de certificados digitales X509 como “Autoridad Certificadora Central” sin limitaciones en el tipo y cantidad de certificados generables.

Utilizando los recursos propios de la plataforma del sistema operativo del nuevo servidor de CENICA, se desarrolló un sistema de autoridad certificadora para la emisión, manejo y revocación de certificados X509. Igualmente se incorporaron al sistema colector y transmisor de datos autónomo todos los módulos para utilizar el protocolo de transferencias SSL (Secure Socket Layers), con autenticación bidireccional por criptografía de llaves públicas y encriptación del canal mediante el uso de los esquemas estándar de más alta seguridad disponibles.

En la actualidad CENICA se encuentra preparando los procedimientos administrativos para la emisión de certificados a todos los participantes del SINAICA.

12. 15 copias del reporte final en español y 15 copias del reporte final en inglés

Esté documento y sus anexos forman el reporte final.

Comentarios y recomendaciones

Haciendo una evaluación general de los resultados del proyecto, en Matías Software Group, creemos haber satisfecho las expectativas generadas.

Sin embargo pensamos que un proyecto de esta naturaleza tiene amplias y variadas posibilidades de crecimiento, en particular sugerimos se considere implementar:

- Los módulos de consulta en múltiples idiomas y múltiples unidades.
- Promedios móviles de 8 y 24 horas.
- Procesos estadísticos adicionales.
- Módulos de apoyo a la validación de datos en las redes.
- Sistemas de información para el público en general, como reportes cartográficos por zona, alertas y recomendaciones para el ejercicio al aire libre, etc.

Así mismo y producto del contacto que nuestros analistas tuvieron con la gente de

Extensión de la cobertura del SINAICA

las redes de monitoreo consideramos que la operación sostenida del SINAICA depende en muy buena medida en que se continúen consolidando tanto los vínculos entre CENICA y las redes, como los apoyos financieros a la operación de las mismas tanto a nivel estatal como federal.

También consideramos importante aumentar el nivel de capacitación tanto operativa como metodológica de todos los involucrados, fundamentalmente en el área de informática.

Es de vital importancia que los procesos normativos se concluyan y consoliden, ésta es probablemente la más sentida de las demandas de las redes.

Por último queremos agradecer de manera especial el apoyo entusiasta que la gente de CENICA nos brindó durante el desarrollo del proyecto, sin él no se habría podido alcanzar las metas propuestas y agradecer a JICA el habernos dado la oportunidad de colaborar al noble fin de mejorar la calidad del aire que respiramos todos.

Atentamente

A nombre de todo el equipo que forma Matías Software Group.
Salvador Ortiz García.
Director General.

México D.F. a 15 de noviembre de 2004.

Anexo 1

Características del tipo y formato de los archivos de datos y tecnología de almacenamiento utilizados en las RAMAs. (Redes Automáticas de Monitoreo Atmosférico)

Introducción.

Este anexo detalla los resultados del proceso de diagnóstico y análisis de las redes incorporadas al SINAICA.

Se procedió a realizar un análisis de la implementación de los sistemas de adquisición, proceso y almacenamiento de datos de las distintas RAMAs así como de la infraestructura de comunicaciones disponible en cada una de ellas, con objeto de determinar las estrategias posibles para la obtención de la información recibida de las estaciones que forman la red y la disponibilidad de medios de comunicación con los equipos de CENICA.

Durante esta etapa del análisis se detectó la siguiente problemática:

Inexistencia de la red como tal.

En varias de la ciudades consideradas y como se anotó en la tabla 1 de los términos de referencia, no se contaba en realidad con una “red de monitoreo”, se cuenta con un conjunto de estaciones de monitoreo automáticas pero no conectadas, en algunos casos con conectividad incipiente hacia alguna de ellas (caso de las ciudades del Bajío), en otros con comunicación hacia sistemas colectores en el extranjero (caso de Ciudad Juárez) y en otros con ninguna capacidad de conectividad (caso de las redes de Baja California).

En todos esos casos, no se encontró un “Centro de Control”, y aunque en algunos casos se contaba con el compromiso por parte de las autoridades locales de establecerlo, éste no fue posible en los plazos del contrato, por lo que se tuvieron que plantear mecanismos alternativos para su incorporación al SINAICA.

Problemas de comunicación hacia el “Centro de Control”

El los casos en que si se encontró un “Centro de Control”, el diagnóstico de los mecanismos de comunicación entre las estaciones que forman la red y el centro de control arrojó lo siguiente:

- Protocolos propietarios en los diferentes dispositivos de adquisición de datos.
- Uso de enlaces seriales RS232 vía “modems”, salvo en los casos de Cd. Juarez, en que se usan enlaces de radio.
- En algunos casos la transmisión se hace mediante líneas dedicadas y en otros usando líneas conmutadas de la red de telefonía pública.

En general los principales problemas que enfrentan en su operación cotidiana las RAMAs están relacionados con esta comunicación, durante el presente contrato, más del 90% de los problemas que reportaron las redes caen es esta clasificación. Aún en las redes operadas por agencias estadounidenses la pérdida de comunicación entre la estación y los sistemas colectores es problema común.

Mientras que por el lado del centro de control de la red se encontró que:

- Algunas redes no cuentan con un sistema de almacenamiento estructurado, obtienen la información de sus estaciones mediante la generación de “reportes” en texto con diversos formatos.
- Cuando se cuenta con almacenamiento estructurado de datos, éste se hace mediante sistemas distintos, cada uno con sus propias definiciones en cuanto a la estructura de los archivos y/o formatos de datos utilizados. El único caso en que se cuenta con un sistema de alto nivel para el almacenamiento de datos es el de la ciudad de Monterrey, en que se utiliza un manejador de bases de datos relacional SQL.
- Generalmente los promedios horarios son calculados localmente.
- En general no se logró determinar la existencia ni de procedimientos establecidos homogéneos ni mecanismos estándar de validación de datos; El conjunto de datos “crudos” recibidos está sujeto en cada red a un proceso manual de validación y de control de calidad, que puede durar de semanas a meses.
- Todos los centros de control de las redes de monitoreo cuentan con algún

grado de conectividad hacia Internet, aunque por las razones expuestas más adelante no se puede considerar que se tenga una interconexión plena.

Deficiencias en el uso de Banderas.

Si bien en todos los casos revisados, los equipos de monitoreo instalados en las estaciones cuentan con la capacidad de asociar a los datos “banderas” que califiquen al dato, únicamente en las estaciones de Ciudad Juárez se hace uso sistemático de esta capacidad. Al respecto se encuentra que:

- Con frecuencia los detectores de diversos estados y condiciones operativas del equipo de monitoreo no se han conectado al “datalogger”.
- A los “dataloggers” no se les ha programado los parámetros de límites de detección, rangos o criterios de abanderamiento.
- Los sistemas utilizados con frecuencia mapean las banderas que indican problemas de detección a algún valor inválido como marcador (p.e. '-99.99'), perdiéndose en ese momento el origen del problema.
- No existen criterios homologados respecto a la semántica y/o codificación de las banderas utilizadas.

El sistema Colector/Transmisor.

En virtud de la problemática detectada, se requiere que el Sistema Colector implemente mecanismos genéricos de importación muy diversos que permitan recuperar los promedios horarios de las mediciones de los medios de almacenamiento nativos de cada red y en algunos casos directamente de los “dataloggers” para poder transmitirlos en forma homologable al sistema concentrador del SINAICA, lo que requiere el desarrollo de un modelo de abstracción de más alto nivel.

Este apartado detalla los mecanismos y la metodología utilizados durante el desarrollo del sistema colector.

En el diseño de los mismos se incorporaron las siguientes consideraciones:

- Minimizar el impacto en la operación de los sistemas existentes.
- Minimizar el impacto en la carga de trabajo de los operadores de la red.
- Maximizar la automatización el proceso de importación de los datos.
- Lograr la mayor generalidad posible.
- Garantizar la seguridad en la transmisión de la información.

Para poder formalizar el término "mecanismos *genéricos*" se requirió elaborar un modelo de abstracción que generalice la noción de 'dato' y que se resume a continuación:

Un 'dato', por ejemplo el promedio horario de una medición, tiene invariablemente asociado un conjunto de 'metadatos' que lo caracterizan:

- Variable medida. (Ozono, NOX, Temperatura, etc.)
- Unidad de medida asociada. (ppm, °C, m/seg, etc.)
- Estación de origen.
- Ubicación temporal. (Fecha y hora)
- Etc.

Dependiendo de la existencia o no de sistemas estructurados y del diseño particular de los mismos usados por una red de monitoreo, los metadatos que caracterizan a un determinado dato serán de una y sólo una de las siguientes clases:

- 1) Constantes.
- 2) Intrínsecos.
- 3) Función del nombre de la tabla o archivo.
- 4) Función del campo en que se localiza.
- 5) Función del valor de otro campo en el registro.

Nótese que:

- Todas las funciones son biunívocas y por tanto con inversa definida.
- Los casos “Reporte de texto” se pueden considerar “tablas” desarrollando los “parsers” necesarios.
- El caso “Datalogger” se puede considerar un caso particular de tabla de acceso secuencial.

La ubicación del dato, dentro de los archivos que formen las tablas del sistema, está determinada en función de aquellos de sus metadatos no constantes o intrínsecos.

De este modelo se infiere que es posible considerar el desarrollo de un "Extractor Universal" que haga abstracción de las diferencias de diseño entre los distintos sistemas utilizados.

Dicho Extractor Universal, mediante simple parametrización, puede obtener los datos utilizando los metadatos como índices.

La parametrización para cada uno de los distintos sistemas de almacenamientos se reduce entonces a clasificar cada uno de los metadatos en alguna de las clases indicadas y a la definición puntual de las funciones de mapeo.

Los diferentes formatos y/o sistemas de almacenamiento de datos encontrados en las redes visitadas y la existencia o no de un “Centro de Control” se resumen en el siguiente cuadro:

Red/Ciudad	Clave	Centro control	Formato o Sistema
México	CMX	SI	DBF
Guadalajara	GDL	SI	DBF
Monterrey	MTY	SI	SQL
Toluca	TLC	SI	DBF
Puebla	PUE	SI	EDAS

Red/Ciudad	Clave	Centro control	Formato o Sistema
CENICA	CEN	Casi	EDAS
Salamanca	SAL	SI	TEXTO
Ciudad Juarez	CJU	NO	TEXTO
Tijuana	TIJ	NO	Datalogger
Mexicali	MXC	NO	Datalogger
Rosarito	ROS	NO	Datalogger
Tecate	TKT	NO	Datalogger

En los Estados Unidos un porcentaje importante de las redes de monitoreo ambiental utilizan en la actualidad el sistema E-DAS de la empresa ESC que almacena la información en formatos propietarios no documentados.

En México, en cambio la mayoría de las redes de monitoreo consideradas utilizan sistemas desarrollados utilizando dBase o herramientas similares, por lo que almacenan la información recibida en bases de datos que utilizan variantes del formato DBF (Data Base Format).

Por ser la clase predominante entre las redes consideradas se tratarán primero las características de este formato.

El formato DBF.

El formato DBF fue desarrollado originalmente por la empresa Ashton Tate para el producto dBASE II en 1983 y su enorme popularidad resultó en varios sistemas "compatibles" entre los que podemos mencionar Clipper y Fox-Pro.

Hoy se le consideran sistemas de base de datos "obsoletos", pero como los desarrollos de los distintos sistemas de almacenamiento utilizados datan de fines de la década de los '90s es comprensible que se le continúe usando en la actualidad.

Los sistemas de bases de datos de las distintas redes de monitoreo que lo utilizan

fueron diseñados y desarrolladas por entidades distintas y no existe una estructura homogénea en cuanto a campos, tipos de datos o nomenclatura de los archivos, lo que impediría el desarrollo de un sistema unificado utilizando cualquiera de las herramientas mencionadas en su forma nativa.

Sin embargo la estructura interna de los archivos DBF en la actualidad está bien documentada, y si bien existen variaciones de la misma producto de las revisiones que el producto dBASE tuvo a lo largo de su ciclo de vida y de las extensiones que le introdujeron las diferentes casas de software que lo adoptaron como formato base, la estructura general del formato interno se mantuvo sin cambios radicales.

Por lo tanto se considera posible obtener directamente del cuerpo de los archivos la información almacenada.

A continuación se detalla la estructura interna de los archivos DBF con el objeto de clarificar la metodología usada por el "Extractor Universal".

En las descripciones siguientes se hace referencia al producto dBASE, pero deberá considerarse en la misma categoría a cualquier otro producto compatible con la versión anotada, salvo en los casos en que existan diferencias que se indiquen explícitamente.

Un archivo en formato DBF está estructurado en tres partes: el encabezado, la descripción de los campos y los registros actuales de datos, tal como se muestra en la figura 1.

Encabezado
Descripción de campos
Registros de datos

Figura 1

El primer byte del encabezado identifica la versión del software utilizado para la creación del archivo, pudiendo ser cualquier de los siguientes:

Versión	Software
02H	dBase II
03H	dBase III-IV
83H	dBase III+ con campos "memo"
8BH	dBase IV+ con campos "memo"
F5H	FoxPro con campos "memo"

Tabla 1

La longitud del encabezado mismo y de los registros de descripción de campos están determinados por este número de versión, en el caso de dBase II (versión 02H), el encabezado ocupa los bytes 0 a 7, seguido por los registro de descripción de campos para los que se reserva espacio para un máximo de 32.

Entonces, en la versión 2 (02H) de un archivo DBF se utilizará la siguiente estructura:

Posición	Longitud	Descripción
0 (00H)	1	Versión (siempre 02h)
1 (01H)	2	Número de registros (0-FFFFH)
3 (03H)	3	Fecha del último acceso de escritura (DDMMYY) En binario
6 (06H)	2	Longitud del registro de datos más uno, hasta 1000.
8-519 (08H-207H)	16*n	16 bytes por cada registro de descripción de campo
16*n+9	1	Marca del fin de encabezado (0DH)

Tabla 2

En esta versión de dBase, como se reserva el espacio para el máximo de 32 campos posibles, los registros de datos comenzarán invariablemente a partir del byte 521 (209H), y el espacio no utilizado en la descripción de campos que siga a la marca

de fin de encabezado aparecerá lleno con el byte cero (00H).

Los 16 bytes que describen cada campo en dBase II tienen la siguiente estructura:

Posición	Longitud	Descripción
0 (00H)	11	Nombre del campo (Cadena ASCII)
11 (0BH)	1	Tipo del campo Ver tabla 6.
12 (0CH)	1	Longitud del campo en bytes Binario 0 hasta FFH
13 (0DH)	2	Reservados
15 (0FH)	1	Posiciones decimales en el campo

Tabla 3

En dBase III, compatibles y versiones posteriores el encabezado tiene siempre una longitud de 32 bytes, seguido de los registros de descripción de campos, en que a diferencia de la versión anterior y habiéndose aumentado el máximo de campos a 128, no se reserva expresamente el espacio para el máximo de campos, los registros de datos seguirán a los registros de descripción de campos.

La tabla siguiente muestra la estructura resultante.

Posición	Longitud	Descripción
0 (00H)	1	Versión (ver tabla 1)
1 (01H)	3	Fecha del último acceso de escritura (YYMMDD) En binario
4 (04H)	4	Número de registros en el archivo 32 bits, sin signo, 'little endian'
8 (08H)	2	Longitud del encabezado 16 bits, sin signo, 'little endian'

Posición	Longitud	Descripción
10 (0AH)	2	Longitud del registro de datos más uno
12 (0CH)	20	Reservados
32 (20H)	32*N	32 bytes por cada registro de descripción de campo
32 * (N+1)	1	Marca de fin del encabezado (ODH)

Tabla 4

En dBase III y posteriores los registros de descripción de campos aumentaron a 32 bytes, con la siguiente estructura:

Posición	Longitud	Descripción
0 (00H)	11	Nombre del campo (Cadena ASCII)
11 (0BH)	1	Tipo del campo Ver tabla 6.
12 (0CH)	4	Reservados
16 (10H)	1	Longitud del campo en bytes
17 (11H)	1	Decimales en el campo
18 (12H)	14	Reservados

Tabla 5

En las descripciones de los campos, el tipo del campo se codifica mediante un único byte, un caracter ASCII de la tabla siguiente, en donde en dBASE II se utilizan únicamente los primeros tres valores.

Caracter	Tipo de campo	Valores posibles
C	Cadena	Caracteres ASCII
N	Numérico	- 0...9

Caracter	Tipo de campo	Valores posibles
L	Booleano	YyNnTtFf
D	Fecha	Numérico, en formato YYYYMMDD
M	Memo	Número de bloque
F	Numérico	- 0...9

Tabla 6

Los registros de datos en todas las versiones de dBase tienen exactamente la misma estructura: una cadena de longitud variable.

Longitud dependiente del número de campos e indicada en el encabezado, en la que el primer byte está reservado como bandera, y los siguientes para la representación en ASCII de los valores del campo sin que se utilice ningún separador o terminador, los campos tipo 'cadena' se completan con espacios en blanco (caracter 32H), así como cualquier campo que no contenga datos.

El primer byte es utilizado por dBASE como un indicador de que el registro ha sido borrado mediante el uso de la operación delete, que por velocidad únicamente lo marca con el carácter '*', posponiéndose hasta el uso de la operación pack el que los registros marcados sean reutilizados y sobrescritos.

Lo anterior debe tomarse en cuenta al leer un registro pues el número de registros que aparece en el encabezado incluye aquellos marcados como borrados.

El offset O de un determinado campo n resulta de la suma de las longitudes L de los campos anteriores.

$$O_n = \sum_{a=1}^{n-1} L_a$$

Cómo se podrá ver, las variaciones en el formato interno de las diferentes versiones se reducen fundamentalmente al tamaño de algunas estructuras y al orden de algunos campos, pero las operaciones involucradas en calcular el índice hacia un

determinado campo dentro de un registro de datos son las mismas.

El "Extractor Universal" utiliza en forma directa los archivos DBF de las distintas bases de datos sin tener que conocer la estructura de alto nivel de la base de datos original.

Las variaciones entre las redes estudiadas.

Ninguna de las tres redes que utilizan este formato usan la misma estructura, aunque para las primeras dos descritas a continuación basta aritmética simple para acceder a los registros:

En el caso de la red de la ciudad de México se usa una tabla mensual por cada variable sujeta a monitoreo, en donde se almacena un registro (renglón) por cada hora, comenzando con el correspondiente a las 0:00 horas del día primero del mes, y en donde los campos (columnas) hacen referencia a las estaciones. Si bien se tiene una columna reservada para el almacenamiento de banderas, en los archivos que la red nos proporciona dicha columna está siempre vacía.

Se considera que ese formato es el más rígido de los usados, pues requiere reestructurar todas las tablas para agregar nuevas estaciones al sistema.

En el caso de la red de la ciudad de Toluca, se usa una tabla por mes por variable y por estación, es en el nombre del archivo mismo donde se codifica la variable y la estación correspondiente, teniéndose simplemente un registro por cada hora y un par de columnas con el valor y la bandera.

Se considera que si bien es el formato más simple es también el más ineficiente por la gran cantidad de archivos involucrados.

En el caso de la red de la ciudad de Guadalajara, se genera un único archivo mensual, en donde se almacena un registro por cada día para cada combinación de variable y estación, y en donde las columnas corresponden a cada una de las horas del día.

Si bien el acceso se vuelve más complejo pues se requiere la creación de un archivo paralelo de índices para evitar hacer búsquedas secuenciales, una vez indexado el archivo resulta en un diseño compacto, eficiente y fácilmente extensible, de ahí que se halla optado por utilizarlo como formato intermedio en los casos en que las redes no contaban con archivos estructurados, caso de Salamanca y Ciudad Juárez y terminar usando el mismo módulo como se describirá más adelante.

El formato de E-DAS de ESC.

El formato interno para el almacenamiento utilizado por el sistema E-DAS es un formato propietario del que Environmental Systems Corporation no proporciona documentación.

El sistema cuenta con mecanismos incorporados de exportación hacia el formato AIRS de la EPA, sin embargo estos no son utilizados en México por lo que no se consideraron como una opción.

A pesar de lo anterior y en virtud de que la legislación aplicable permite el análisis de formatos propietarios con objeto de interoperabilidad con otros sistemas se utilizaron técnicas de decodificación que permitieron la incorporación de este formato al modelo del "Extractor Universal".

A continuación se presentan los detalles del formato:

Por cada mes se tiene un archivo de datos, el nombre indica el mes y el año de la lectura. Lleva tres letras seguidas de cuatro dígitos y la extensión ".dat". Los dos primeros dígitos se refieren a las dos cifras menos significativas del año y las dos siguientes al mes. Así, el archivo correspondiente a mayo del 2003 se llama "hly0305.dat."

El archivo contiene exclusivamente registros, sin encabezado. El tamaño de los registros es de 648 bytes y son de posición fija. Se tiene un registro por cada día, estación y medida. Dentro de cada registro ocurren primero un grupo de campos que describen el registro y a continuación los valores del día.

En la tabla se describen las posiciones y los campos para el primer grupo:

Posición	Longitud	Descripción
0 (00H)	2	Número de estación
2 (02H)	2	Número de variable
4 (04H)	8	Nombre de la variable
12 (0CH)	8	Unidad de medida de la variable
20 (14H)	8	Nombre de la estación
28 (1BH)	92	Valores constantes

A partir de la posición 120 se encuentran los datos por hora, 24 por registro, es decir uno por hora. La longitud de cada grupo es de 22 caracteres.

En la tabla se muestran las posiciones, campos y descripciones para cada grupo, con posición relativa al grupo:

Posición	Longitud	Descripción
0 (00H)	4	Variable de tipo punto flotante con el valor de medida de la variable
4 (04H)	6	Banderas de la lectura
10 (0AH)	1	Espacio en blanco
11 (0BH)	1	Serial con la hora
12 (0CH)	9	Campo reservado

Por ejemplo:

00000000:	3031	3031	5753	2020	2020	2020	4d2f	5320	0101WS	M/S
00000010:	2020	2020	5055	4542	4c41	2d31	0100	0100	PUEBLA-1
00000020:	9af9	79c4	9a3f	1c46	2a2a	2a2a	2a2a	2a2a	..y..?.F	*****
00000030:	2a2a	*****	*****							
00000040:	2a2a	*****	*****							
00000050:	2a2a	2a2a	2a2a	2020	2020	2020	2020	202a	*****	*
00000060:	2a2a	*****	*****							
00000070:	2a2a	2a2a	2a2a	2a2a	ed64	0640	0000	0000	*****	.d.@....
00000080:	0000	2000	0000	0000	0000	0020	202a	2dbd *-.
00000090:	8940	0000	0000	0000	2001	0000	0000	0000	..@.....
000000a0:	0020	202a	0590	9a40	0000	0000	0000	2002	..*	...@.....
000000b0:	0000	0000	0000	0020	202a	6dd0	d440	0000	*m..@..
000000c0:	0000	0000	2003	0000	0000	0000	0020	202a*
000000d0:	4900	c640	0000	0000	0000	2004	0000	0000	I..@.....
000000e0:	0000	0020	202a	f84b	c840	0000	0000	0000	...*	.K.@.....
000000f0:	2005	0000	0000	0000	0020	202a	4b8c	a040	*K..@
00000100:	0000	0000	0000	2006	0000	0000	0000	0020
00000110:	202a	282f	ab40	0000	0000	0000	2007	0000	*(/. @.....	...
00000120:	0000	0000	0020	202a	e4a9	5640	0000	0000	*.V@.....
00000130:	0000	2008	0000	0000	0000	0020	202a	e92c*,
00000140:	0c40	0000	0000	0000	2009	0000	0000	0000	..@.....
00000150:	0020	202a	271c	6440	0000	0000	0000	200a	..*	'd@.....
00000160:	0000	0000	0000	0020	202a	1352	aa40	0000	*.R.@..
00000170:	0000	0000	200b	0000	0000	0000	0020	202a*
00000180:	651f	8d40	0000	0000	0000	200c	0000	0000	e..@.....
00000190:	0000	0020	202a	1410	b040	0000	0000	0000	...*	...@.....
000001a0:	200d	0000	0000	0000	0020	202a	b4cf	8740	*...@
000001b0:	0000	0000	0000	200e	0000	0000	0000	0020
000001c0:	202a	c15d	8a40	0000	0000	0000	200f	0000	*.].@.....	...
000001d0:	0000	0000	0020	202a	0d58	a240	0000	0000	*.X.@....
000001e0:	0000	2010	0000	0000	0000	0020	202a	5d3b*];
000001f0:	9440	0000	0000	0000	2011	0000	0000	0000	..@.....
00000200:	0020	202a	e39c	0e40	0000	0000	0000	2012	..*	...@.....
00000210:	0000	0000	0000	0020	202a	f5e3	0d40	0000	*...@..
00000220:	0000	0000	2013	0000	0000	0000	0020	202a*
00000230:	e509	d93f	0000	0000	0000	2014	0000	0000	...?
00000240:	0000	0020	202a	f500	8040	0000	0000	0000	...*	...@.....
00000250:	2015	0000	0000	0000	0020	202a	AAF0	7340	*..s@
00000260:	0000	0000	0000	2016	0000	0000	0000	0020
00000270:	202a	efdc	5c40	0000	0000	0000	2017	0000	*.. \	@.....
00000280:	0000	0000	0020	202a	3031	3032	5744	2020	*0102WD
00000290:	2020	2020	4752	4144	4f53	2020	5055	4542	GRADOS	PUEB
000002a0:	4c41	2d31	0100	0100	9af9	79c4	9a3f	1c46	LA-1y..?.F
000002b0:	2a2a	*****	*****							

Podemos ver un registro completo y el primer grupo del siguiente.

El encabezado se puede ver en la figura 1:

```

00000000: 3031 3031 5753 2020 2020 2020 4d2f 5320 0101WS M/S
00000010: 2020 2020 5055 4542 4c41 2d31 0100 0100 PUEBLA-1...
00000020: 9af9 79c4 9a3f 1c46 2a2a 2a2a 2a2a 2a2a ..y..?.F*****
00000030: 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a *****
00000040: 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a *****
00000050: 2a2a 2a2a 2a2a 2020 2020 2020 2020 202a ***** *
00000060: 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a 2a2a *****
00000070: 2a2a 2a2a 2a2a 2a2a *****
    
```

Figura 1 Encabezado de un registro

```

00000070: ed64 0640 0000 0000 .d.@...
00000080: 0000 2000 0000 0000 0000 0020 202a .. ..... *
    
```

Figura 2 Primer valor del registro

En la figura 2 se ve el valor, que es el primero del registro. Como se observa los primeros cuatro bytes tienen el valor de la medida. Aplastando el subregistro e iniciando la cuenta en cero:

```

ed 64 06 40 00 00 00 00 00 00 20 00 00 00 00 00 00 00 20 20 2a
 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21
    
```

El valor es un *float* nativo de Intel y es *little-endian*. En el byte número once se almacena el serial de la hora. Comienza en cero, pero quizá tiene un desplazamiento.

Del byte número cuatro al nueve están las banderas. En éste ejemplo, las banderas están en ceros.

```

00003350:                                acef b340 4010 ..... *...@@.
00003360: 0000 0000 2003 0000 0000 0000 0020 202a .... ..... *
    
```

Figura 3 Registro con valores en la Bandera

En este otro ejemplo, correspondiente al registro mostrado en la figura 3, el valor de la bandera correspondiente es: 401000000000 en hexadecimal. Observemos, también, que el serial de la hora vale tres en la posición once.

```

ac ef b3 40 40 10 00 00 00 00 20 03 00 00 00 00 00 00 20 20 2a
 0  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21
    
```

Archivos de índices

Para cada archivo de datos hay un archivo de índices. Tiene el mismo nombre pero con extensión “.inx”. Por ejemplo, el archivo de índices correspondiente a **hly0307.dat** es **hly0307.inx**.

Este archivo también está compuesto por registros de tamaño fijo, de 128 bytes. Cada registro se compone de un pequeño encabezado de veinte bytes en los cuales aparece los datos de la estación y la variable. En los dos primeros bytes aparece la variable y en los dos siguientes la variable. El resto del encabezado contiene la constante '001H*****'.

En cada registro, a continuación del encabezado, ocurren cincuenta y cuatro campos de enteros cortos. Los primeros treinta y un campos coinciden con los días del calendario, los siguientes veintitrés aparecen en ceros. Los meses que tienen menos de treinta y un días, tienen ceros en los campos que no utilizan.

En el archivo de índices hay un registro por estación y variable. En cada registro hay una lectura por cada día en uno de los primeros treinta y un campos. En el caso de que no se haya hecho la lectura, aparece un cero. En caso contrario aparece el número del registro en el archivo de datos que contiene los datos de la lectura en sí. Esto es, que el procedimiento para leer un registro del archivo de datos dados la estación y la variable a leer, pasa por localizar en el índice el registro con dicha pareja y posicionarse dentro del registro en el día deseado. El valor leído es el número de registro que se deberá leer del archivo de datos.

Descripción del campo de banderas.

Las banderas posibles que ocurren en cada uno de los campos del registro, están divididas en banderas de datos y banderas de estado. En la tabla 1 se presentan las primeras y en la tabla 2 las de estado.

De acuerdo a los listados que examinamos, las seis banderas utilizadas son '<', 'D', 'C', 'M', '-', 'V' e 'I'.

De los seis bytes de banderas, en los primeros tres se representan las banderas de datos y en los últimos tres las de estado. Las banderas más empleadas y la máscara con la que se obtienen se puede ver en la tabla 3, tomando exclusivamente los tres primeros bytes.

En el caso de las banderas de estado, la bandera que indica que no se debe de utilizar la lectura se obtiene con la máscara 0x000100, sobre los tres bytes menos significativos. En ningún caso encontramos otra bandera prendida en éste grupo.

Bandera	Significado
<	Less than ##% Data
P	Power Fail
D	Disabled
T	Out-of-Control
F	Boiler Off-Line
B	Bad Status
C	Calibration
M	Maintenance
O	Analog Over range
U	Analog Under range
A	Arithmetic Error
+	Maximum
-	Minimum
R	Rate of Change
H	High-High Alarm
L	Low-Low Alarm
h	High Alarm
l	Low Alarm
J	High Rate of Change

Bandera	Significado
j	Low Rate of Change
V	DIS #1 Obs
W	DIS #2 Obs
X	DIS #3 Obs
Y	DIS #4 Obs
Z	DIS #5 Obs.
f	Sin documentación
	<i>Cuadro 1 Banderas de datos</i>

Bandera	Significado
I	Ignorar?
?	Sin documentar
□	Sin documentar
>	Sin documentar
=	Sin documentar
m	Sin documentar
^	Sin documentar
v	Sin documentar
E	Sin documentar
d	Sin documentar
9	Sin documentar
z	Sin documentar
a	Sin documentar
Q	Sin documentar

Bandera	Máscara
<	0x001000
D	0x004000
C	0x040000
M	0x080000
-	0x000100
V	0x000001
	<i>Cuadro 3. Banderas de datos</i>

El formato de la RDBMS en SQL

El uso de un Manejador de Bases de Datos Relacionales que además utilice un lenguaje de alto nivel como SQL para el almacenamiento simplifica enormemente la extracción de los datos.

En el caso de la red de la ciudad de Monterrey se tiene instalado MS-SQL Server (Servidor SQL de Microsoft) que si bien es un producto comercial con costos elevados de licenciamiento en realidad no es más que una versión especializada para el sistema operativo Microsoft Windows NT del manejador desarrollado por Sybase, Inc. del que se cuenta con amplia documentación y soporte en diversos sistemas operativos.

Utilizando las bibliotecas estándar del lenguaje Perl5 para Sybase SQL Server, que no tienen costo de licenciamiento, la extracción de los datos se redujo a una sentencia simple en SQL:

```
SELECT * FROM hravedata
WHERE parameter = ? AND date >= ? AND date <= ?
ORDER BY date;
```

Cabe mencionar que la implementación del sistema en Monterrey utiliza una base

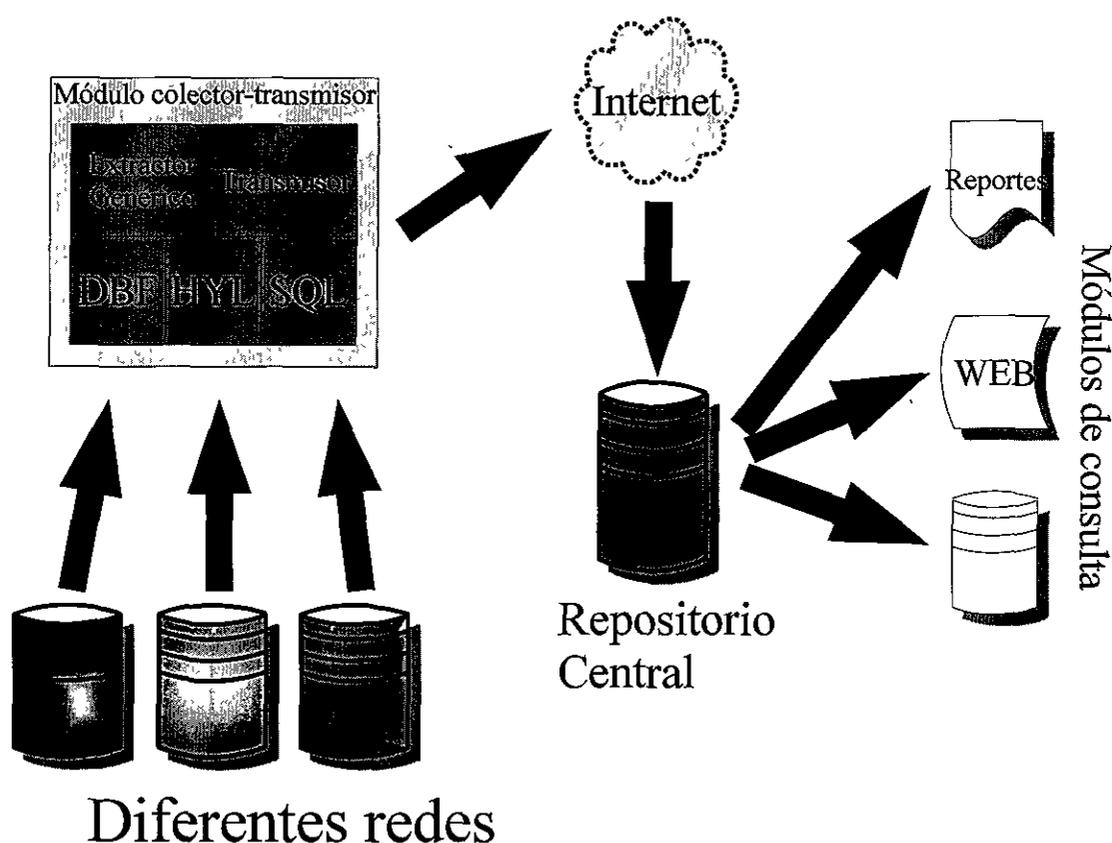
de datos separada para cada estación, bases de datos con nombres “SUROESTE”, “NOROESTE”, “CENTRO”, “NOESTE” y “SURESTE” por lo que se requiere interrogar separadamente a cada una de ellas.

Anexo 2 Implementación

Introducción.

Este anexo describe los detalles de la implementación de la nueva versión del SINAICA.

A continuación muestra en forma esquemática el el flujo de datos entre los diferentes componentes del SINAICA:



Es importante hacer notar que éste esquema general de operación ha demostrado ser muy exitoso debido fundamentalmente a su modularidad, que lo vuelve un esquema flexible y extensible.

De forma que esta nueva versión del SINAICA de forma alguna pretendió alterar ese esquema general, se buscó al mismo tiempo que se extendía la cobertura del

sistema a nuevas redes, robustecerlo y enriquecer sus capacidades.

Este contrato nos dio la oportunidad de reimplementar algunos componentes, extender la funcionalidad de otros y reducir la dependencia tecnológica del sistema en general.

Sistema Colector y Transmisor de Dato autónomo.

En versiones anteriores del SINAICA, la extracción de los datos de los sistemas propios de cada red y su transmisión hacia la base de datos central estaba delegado un programa normalmente bajo el sistema operativo MS Windows, que correría en alguna de las computadoras del centro de control.

Este esquema presentó múltiples inconvenientes detectados en las versiones anteriores, inconvenientes que pueden resumirse en un aumento significativo tanto en los costos de operación como en la carga de trabajo de los operadores de las redes.

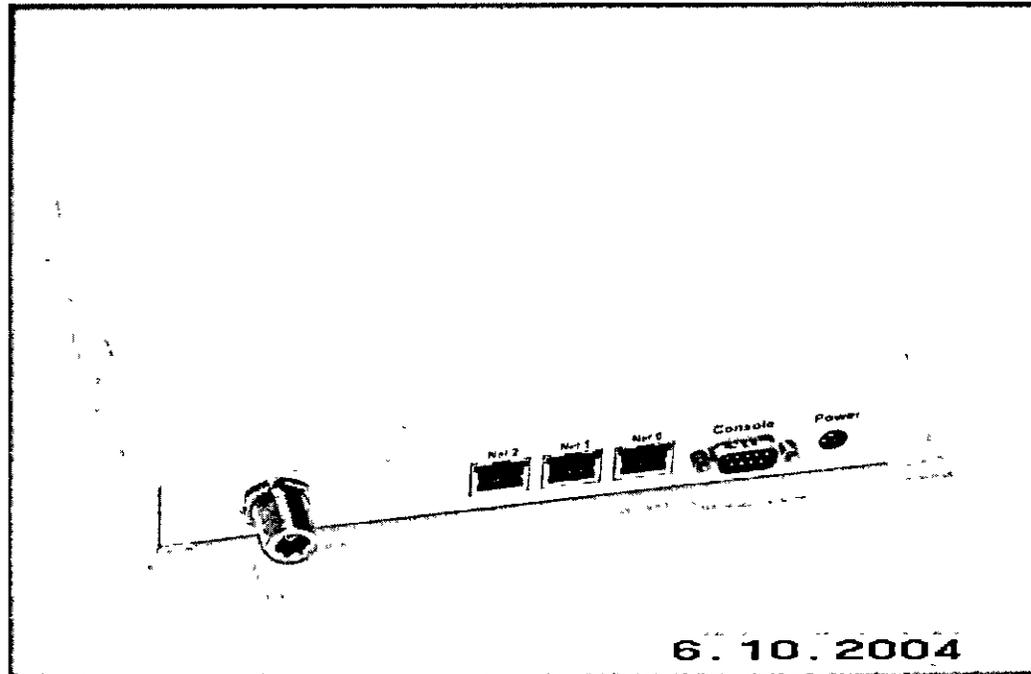
Se partió de la base de que el nuevo sistema tenía que cumplir con ocho características esenciales: robustez, eficiencia, extensibilidad, portabilidad y facilidad de soporte, al mismo tiempo que redujera los costos asociados.

Para cumplir con los lineamientos propuestos se estableció como primer requisito que todo el software utilizado debería ser abierto y sin costos de licenciamiento, dado que esto garantizaría la portabilidad y facilidad de soporte además de incidir en el aspecto económico; En el Anexo 3 se mencionan otras características del software abierto “Software Libre”.

Matías Software Group había desarrollado para otros proyectos de comunicación la plataforma “MSG Communication Server” utilizando una variante especial del sistema operativo Linux para sistemas incrustados. El hardware utilizado es una computadora embebida basada en el procesador Elan520, arquitectura compatible con Intel, de bajo costo, muy bajo consumo de energía y cero mantenimiento, que no utiliza piezas móviles y es fácilmente reemplazable.

Consideramos que los módulos extractores podrían adecuarse y portarse a esta plataforma, eliminando la dependencia de una computadora personal dedicada, de estos esfuerzos surgió el Sistema Colector y Transmisor de Datos (SCTD) autónomo.

La fotografía a continuación muestra el aspecto del SCTD autónomo:



Administración remota

Para atacar este punto se integró un módulo que permite consultar las bitácoras de los SCTD en tiempo real mediante una página Web accesible en la intranet de las redes de monitoreo. Dicho módulo permite el diagnóstico oportuno de problemas y coadyuva al soporte técnico centralizado. Así mismo, como parte del módulo de consulte especializada se integró un sistema que lleva una bitácora centralizada de las transferencias por red, así como el reporte automatizado de efemérides y eventos relevantes.

Anexo 3

Detalles técnicos

Introducción.

Este documento detalla las características técnicas de los diferentes programas y subsistemas que componen la tercera versión del SINAICA.

Plataformas.

Los diferentes módulos del sistema corren en una o mas de tres plataformas distintas:

- Fedora Core 2 (Kernel Linux v2.6.x) para i386 [FC2]
- MSG Communication Server (Kernel Linux v2.4.x) para ELAN520 [MSC]
- MS-Windows 98 svr2 o superior para i386 [WIN]

En la primera de las mencionadas corren el manejador de base de datos y el servidor HTTP que conjuntamente constituyen el subsistema de almacenamiento central y de consultas.

En la segunda corren los sistemas colector/extractor y transmisor que se instalan en las redes de monitoreo.

Por último, en MS Windows corren algunos de los módulos de manejo de datos manuales, como el sistema de captura para globos piloto o el sistema de manejo de carbón orgánico y elemental.

Requerimientos.

Toda la funcionalidad del sistema descansa en el uso de cuatro herramientas fundamentales: El sistema operativo Linux, el manejador de bases de datos relacional PostgreSQL, el servidor de HTTP GranJefe (derivado de Apache) y el lenguaje de programación Perl5, todas de ente los mejores frutos del movimiento internacional del “software libre” y por supuesto libres de costos de licenciamiento y de libre distribución.

Si bien el núcleo del sistema está compuesto por software libre, se ha cuidado que ningún componente del sistema requiera, por marginal que sea, del uso de software comercial sujeto a costos de licenciamiento. Incluso los componentes desarrollados para ser ejecutados en la plataforma Windows de Microsoft se han programado usando técnicas que permiten su operación en otras plataformas y no requieren de componentes exclusivos de ese sistema operativo.

Lo anterior buscando desde las etapas iniciales de análisis y diseño el que todos los componentes se puedan distribuir libremente y la comunidad nacional e internacional pueda aprovechar estos desarrollos sin incurrir en costos de licenciamiento.

El único módulo que se aparta de estos lineamientos es el paquete de macros de Excel desarrollados para el proceso estadístico de los datos de COVs, por así requerirlo explícitamente CENICA en virtud de que se tiene capturado una gran cantidad de datos en ese sistema.

Parámetros de compilación.

El desarrollo de los componentes del sistema se realizó utilizando el lenguajes de programación Perl5.

Perl5 es un lenguaje de programación interpretado por lo que los módulos ejecutables no requieren del proceso de compilación.

En todos los casos se utilizó el interprete estándar de la plataforma involucrada y tampoco se requiere compilarlo, por completes, se anota a continuación los parámetros con que se compiló originalmente en las plataformas utilizadas.

Perl en Linux: (Perl 5.8.3 en Fedora Core 2)

```
Summary of my perl5 (revision 5.0 version 8 subversion 3) configuration:
Platform:
  osname=linux, osvers=2.4.21-4.el5mp, archname=i386-linux-thread-multi
  uname='linux tweety.devel.redhat.com 2.4.21-4.el5mp #1 smp fri oct 3 17:52:56
  edt 2003 i686 i686 i386 gnulinux '
  config_args='-des -Doptimize=-O2 -g -pipe -march=i386 -mcpu=i686
-Dversion=5.8.3 -Dmyhostname=localhost -Dperladmin=root@localhost -Dcc=gcc
-Dcf_by=Red Hat, Inc. -Dinstallprefix=/usr -Dprefix=/usr -Darchname=i386-linux
-Dvendorprefix=/usr -Dsiteprefix=/usr -Duseshrplib -Duseithreads -Duseithreads
```

```

-Duselargefiles -Dd_dosuid -Dd_semctl_semun -Di_db -Ui_ndbm -Di_gdbm -Di_shadow
-Di_syslog -Dman3ext=3pm -Duseperl10 -Dinstallusrbinperl -Ubincompat5005
-Uversiononly -Dpager=/usr/bin/less -isr -Dinc_version_list=5.8.2 5.8.1 5.8.0'
    hint=recommended, useposix=true, d_sigaction=define
        usethreads=define      use5005threads=undef      useithreads=define
usemultiplicity=define
    useperl10=define d_sfio=undef uselargefiles=define usesocks=undef
    use64bitint=undef use64bitall=undef uselongdouble=undef
    usemymalloc=n, bincompat5005=undef
Compiler:
    cc='gcc', ccflags ='-D_REENTRANT -D_GNU_SOURCE -DTHREADS_HAVE_PIDS -DDEBUGGING
-fno-strict-aliasing -I/usr/local/include -D_LARGEFILE_SOURCE
-D_FILE_OFFSET_BITS=64 -I/usr/include/gdbm',
    optimize='-O2 -g -pipe -march=i386 -mcpu=i686',
    cppflags='-D_REENTRANT -D_GNU_SOURCE -DTHREADS_HAVE_PIDS -DDEBUGGING -fno-
strict-aliasing -I/usr/local/include -I/usr/include/gdbm'
    ccversion='', gccversion='3.3.3 20040412 (Red Hat Linux 3.3.3-7)',
gccosandvers=''
    intsize=4, longsize=4, ptrsize=4, doublesize=8, byteorder=1234
    d_longlong=define, longlongsize=8, d_longdbl=define, longdblsize=12
    ivtype='long', ivsize=4, nvtype='double', nvsize=8, Off_t='off_t', lseeksize=8
    alignbytes=4, prototype=define
Linker and Libraries:
    ld='gcc', ldflags = ' -L/usr/local/lib'
    libpth=/usr/local/lib /lib /usr/lib
    libs=-lnsl -lgdbm -ldb -ldl -lm -lcrypt -lutil -lpthread -lc
    perllibs=-lnsl -ldl -lm -lcrypt -lutil -lpthread -lc
    libc=/lib/libc-2.3.3.so, so=so, useshrplib=true, libperl=libperl.so
    gnulibc_version='2.3.3'
Dynamic Linking:
    dlsrc=dl_dlopen.xs, dlexit=so, d_dlsymun=undef, ccdlflags='-rdynamic -Wl,-
rpath,/usr/lib/perl5/5.8.3/i386-linux-thread-multi/CORE'
    cccdflags='-fPIC', lddlflags='-shared -L/usr/local/lib'
Compiled at Apr 15 2004 13:09:17

```

Perl en MSG Communications Server

Summary of my perl5 (revision 5 version 8 subversion 4) configuration:

```

Platform:
    osname=linux, osvers=2.4.26-lfs-i386-fc1, archname=i386-linux-thread-multi
    uname='linux xochitl 2.4.26-lfs-i386-fc1 #3 thu jul 15 21:10:51 cdt 2004 1686
genuineintel unknown gnulinux '
    config_args=''
    hint=previous, useposix=true, d_sigaction=define
        usethreads=define      use5005threads=undef      useithreads=define
usemultiplicity=define
    useperl10=define d_sfio=undef uselargefiles=undef usesocks=undef
    use64bitint=undef use64bitall=undef uselongdouble=undef
    usemymalloc=n, bincompat5005=undef
Compiler:
    cc='cc', ccflags ='-D_REENTRANT -D_GNU_SOURCE -DTHREADS_HAVE_PIDS -fno-strict-
aliasing -I/usr/include/db-1.86',
    optimize='-O2',
    cppflags='-D_REENTRANT -D_GNU_SOURCE -DTHREADS_HAVE_PIDS -fno-strict-aliasing
-I/usr/include/db-1.86 -D_REENTRANT -D_GNU_SOURCE -DTHREADS_HAVE_PIDS -fno-strict-
aliasing -I/usr/include/db-1.86'
    ccversion='', gccversion='3.3.3', gccosandvers=''
    intsize=4, longsize=4, ptrsize=4, doublesize=8, byteorder=1234

```

```

d_longlong=define, longlongsize=8, d_longdbl=define, longdblsize=12
ivtype='long', ivsize=4, nvtype='double', nvsize=8, Off_t='off_t', lseeksize=4
alignbytes=4, prototype=define
Linker and Libraries:
ld='cc', ldflags=''
libpth=/lib /usr/lib
libs=-lnsl -ldb -ldl -lm -lcrypt -lutil -lpthread -lc
perllibs=-lnsl -ldl -lm -lcrypt -lutil -lpthread -lc
libc=/lib/libc-2.3.3.so, so=so, useshrplib=true, libperl=libperl.so
gnulibc version='2.3.3'
Dynamic Linking:
dlsrc=dl_dlopen.xs, dlext=so, d_dlsymun=undef, ccdlflags='-Wl,-E -Wl,-
rpath,/usr/lib/perl5/5.8.4/i386-linux-thread-multi/CORE'
cccdlflags='-fpic', lddlflags='-shared'
Compiled at Jul 16 2004 16:50:08

```

Perl en Windows: (Perl 5.8.0 build 806 de ActiveState Corp.)

Summary of my perl5 (revision 5 version 8 subversion 4) configuration:

```

Platform:
osname=MSWin32, osvers=4.0, archname=MSWin32-x86-multi-thread
uname=''
config_args='undef'
hint=recommended, useposix=true, d_sigaction=undef
usethreads=undef use5005threads=undef useithreads=define usemultiplicity=define
useperlio=define d_sfio=undef uselargefiles=define usesocks=undef
use64bitint=undef use64bitall=undef uselongdouble=undef
usemymalloc=n, bincompat5005=undef
Compiler:
cc='cl', ccflags='-nologo -Gf -W3 -MD -Zi -DNDEBUG -O1 -DWIN32 -D_CONSOLE
-DNO_STRICT -DHAVE_DES_FCRYPT -DNO_HASH_SEED -DPERL_IMPLICIT_CONTEXT
-DPERL_IMPLICIT_SYS -DUSE_PERLIO -DPERL_MSVCRT_READFIX',
optimize='-MD -Zi -DNDEBUG -O1',
cppflags='-DWIN32'
ccversion='', gccversion='', gccosandvers=''
intsize=4, longsize=4, ptrsize=4, doublesize=8, byteorder=1234
d_longlong=undef, longlongsize=8, d_longdbl=define, longdblsize=10
ivtype='long', ivsize=4, nvtype='double', nvsize=8, Off_t='__int64', lseeksize=8
alignbytes=8, prototype=define
Linker and Libraries:
ld='link', ldflags='-nologo -nodefaultlib -debug -opt:ref,icf -libpath:
"C:\Perl\lib\CORE" -machine:x86'
libpth=C:\PROGRA~1\MICROS~3\VC98\lib
libs= oldnames.lib kernel32.lib user32.lib gdi32.lib winspool.lib comdlg32.lib
advapi32.lib shell32.lib ole32.lib oleaut32.lib netapi32.lib uuid.lib wsock32.lib
mpr.lib winmm.lib version.lib odbccp32.lib msvcrt.lib
perllibs= oldnames.lib kernel32.lib user32.lib gdi32.lib winspool.lib
comdlg32.lib advapi32.lib shell32.lib ole32.lib oleaut32.lib netapi32.lib uuid.lib
wsock32.lib mpr.lib winmm.lib version.lib odbccp32.lib msvcrt.lib
libc=msvcrt.lib, so=dll, useshrplib=yes, libperl=perl58.lib
gnulibc_version='undef'
Dynamic Linking:
dlsrc=dl_win32.xs, dlext=dll, d_dlsymun=undef, ccdlflags=' '
cccdlflags=' ', lddlflags='-dll -nologo -nodefaultlib -debug -opt:ref,icf
-libpath:"C:\Perl\lib\CORE" -machine:x86'

Compiled at Jun 1 2004 11:52:21

```

Relación de programas desarrollados y bibliotecas utilizadas.

La siguiente tabla es la relación completa de los programas desarrollados y las bibliotecas que cada uno de ellos utiliza.

Programa	Función	Plataforma
td3gui.pl	Programa de captura para teodolitos TD3.	WIN
	Win32API::CommPort Win32::SerialPort Win32::API Win32::GUI DBI (1.43) Globo::Query Globo::Graph Globo::Calc Math::Trig (1.02) Globo::Parse Globo::TD3 Time::Local (1.07) Tk (804.027) Win32::TieRegistry (0.25)	
carbongui.pl	Programa de captura COR y CEL	WIN
	Win32::API Win32::GUI DBI (1.43) DBD::ODBC Tk (804.027) Win32::TieRegistry (0.25)	
rmaxtr	Sistema colector/extractor y transmisor	MCS

	RMA::Storage MIME::Base64 (2.12) Storable (2.06) RMA::Transfer LWP::UserAgent (2.001) Symbol (1.04) RMA::Xtract Time::Local (1.04) RMA::Xtract::DataLogger Device::SerialPort (1.000002) RMA::Xtract::DBF XBase (0.241) RMA::Xtract::HLY RMA::Xtract::Mzip Sys::Syslog (0.03)	
dat2dbf	<i>Parser</i> del formato de texto usado en Salamanca.	MCS
epa2dbf	<i>Parser</i> del formato de texto usado por Cd. Juarez	MCS
	DB_File (1.808) Fcntl (1.05) File::Basename (2.72) Time::Local (1.07) XBase::Fast	
RMA::PcdCA.pm	Servicio GJ de recepción de datos:	FC2
	Data::Dumper (2.121) MIME::Base64 (2.21) RMA::Saver DBI (1.42) DBD::Pg (1.32) RMA::Normalizer Time::Local (1.07) Storable (2.09) Unicode::String (2.07) XML::Simple (2.12)	
MSG::RMA.pm	Servicio GJ de consultas	FC2

DBI (1.42)
DBD::Pg (1.32)
Encode (1.99)
MSG::RMA::Graf
SVG::TT::Graph::Bar (0.06-msg)
SVG::TT::Graph::Line (0.06)
SVG::TT::Graph::TimeSeries (0.06-msg)
MSG::RMA::XRF
Time::Local (1.07)

Anexo 4

Licencia

El texto a continuación corresponde a una versión en español de la GPL, y constituye la base y el espíritu de la propuesta de licencia con que Matías Software Group pretende licenciar sus desarrollos relativos al SINAICA, pero no necesariamente es la licencia definitiva pues se están llevando a cabo pláticas con el INE para encontrar los términos precisos que mejor representen los intereses de todas las partes y mejor se adecuen a la legislación mexicana.

Las notas siguiente aparece en cumplimiento del requisito que la Free Software Foundation impone para la licencia GPL en cualquier traducción a idiomas distintos del inglés.

NOTA IMPORTANTE:

Esta es una traducción no oficial al español de la GNU General Public License. No ha sido publicada por la Free Software Foundation, y no establece legalmente las condiciones de distribución para el software que usa la GNU GPL. Estas condiciones se establecen solamente por el texto original, en inglés, de la GNU GPL. Sin embargo, esperamos que esta traducción ayude a los hispanoparlantes a entender mejor la GNU GPL.

IMPORTANT NOTICE:

This is an unofficial translation of the GNU General Public License into Spanish. It was not published by the Free Software Foundation, and does not legally state the distribution terms for software that uses the GNU GPL--only the original English text of the GNU GPL does that. However, we hope that this translation will help Spanish speakers understand the GNU GPL better.

Copyright (C) 1989, 1991 Free Software Foundation, Inc.
675 Mass Ave, Cambridge, MA 02139, EEUU

LICENCIA PÚBLICA GENERAL GNU

Versión 2, Junio de 1991

Copyright © 1989, 1991 Free Software Foundation, Inc.

59 Temple Place – Suite 330, Boston, MA 02111-1307, USA

Cualquier persona tiene el permiso de copiar y distribuir copias fieles de esta licencia, sin embargo, no está permitido modificarla en ninguna forma.

Preámbulo

Las licencias de la mayoría de los programas de cómputo están diseñadas para coartar la libertad de compartirlos y cambiarlos. Por el contrario, la Licencia Pública General GNU pretende garantizar esa libertad de compartir y cambiar Software Libre a fin de asegurar que el software sea libre para todos sus usuarios. Esta Licencia Pública General se aplica a la mayor parte del software de la Free Software Foundation y a cualquier otro programa cuyos autores se comprometan a usarla. (Algunos otros paquetes de software de la Free Software Foundation están protegidos bajo la Licencia Pública General de Librería GNU.) Esta última licencia también puede aplicarse a nuevos paquetes de software.

Cuando se hable de Software Libre, se hace referencia a libertad, no a precio. Las Licencias Públicas Generales GNU están diseñadas para asegurar que el usuario tenga libertad de distribuir copias de Software Libre (y de recibir una remuneración por este servicio, si así se desea), que ese mismo usuario reciba el código fuente o que tenga la posibilidad de recibirlo, si así lo desea, que pueda cambiar o modificar el software o utilice sólo partes del mismo en nuevos paquetes de Software Libre; y que dicho usuario tenga pleno conocimiento de estas facultades.

Con la finalidad de proteger los derechos antes mencionados, es necesario establecer restricciones que prohíban a cualquiera negar esos derechos o pedir la renuncia a los mismos. Estas restricciones se traducen en ciertas responsabilidades para el usuario que distribuye o modifica copias de software protegido bajo estas licencias.

Por ejemplo, si una persona distribuye copias de un paquete de Software Libre protegido bajo esta licencia, ya sea de manera gratuita o a cambio de una contraprestación, esa persona debe dar a los receptores de esa distribución todos y cada uno de los derechos que él o ella misma tenga. Asimismo, esa persona debe asegurarse que dichos receptores reciban o tengan la posibilidad de recibir el código fuente. De igual manera, debe mostrarles esta licencia a fin de que tengan conocimiento de los derechos de los que son titulares.

La protección que otorga la presente licencia se hace de dos maneras simultáneas: (1) se otorga protección al software bajo la ley de copyright, y (2) se ofrece la protección bajo esta licencia, la cual otorga permiso legal para copiar, distribuir y/o modificar el software.

Asimismo, a fin de proteger a cada uno de los autores y a los creadores mismos de esta licencia, es importante hacer notar y que todos entiendan que no existe

ninguna garantía de cualquier paquete de Software Libre por la cual se deba responder. Esto es, si el software es modificado por alguna persona distinta del autor y distribuido con esas modificaciones, los receptores de esa distribución deben saber que lo que han recibido no es la obra original, y que por lo tanto, cualquier problema surgido de las modificaciones no se reflejará en la reputación del autor original.

Finalmente, cualquier programa de Software Libre es amenazado por patentes de Software. Esta licencia tiene la finalidad de evitar el peligro que representa que los redistribuidores de programas de Software Libre obtengan individualmente licencias de patentes, haciendo de esta forma, programas de Software Propietario. Para lograr esto, queda totalmente claro que cualquier patente debe otorgar licencias que permitan el uso libre del programa para todos o no otorgar licencia alguna.

Los términos y condiciones específicos para copiar, distribuir o modificar son los siguientes:

TÉRMINOS Y CONDICIONES PARA LA COPIA, DISTRIBUCIÓN Y MODIFICACIÓN

0. Esta licencia se aplica a cualquier programa u otra obra que contenga un aviso puesto por el titular de los derechos de autor en el que se establezca que el mismo puede ser distribuido bajo los términos de esta Licencia Pública General. El “Programa” se refiere a cualquier programa u obra, y “Obra basada en el Programa” se refiere por su parte, a, ya sea al “Programa” mismo a cualquier obra derivada del mismo según la ley de Derechos de Autor; esto es, una obra que contenga el “Programa” o una porción del mismo, ya sea que esta porción sea exactamente igual o modificada y/o traducida a otro idioma. (En adelante, una traducción se considerará de manera enunciativa, mas no limitativa, como una “modificación”.)

Actividades distintas de copiar o distribuir no son abarcadas por esta licencia; están fuera de su alcance. El acto de correr el “Programa” no está restringido, y el producto que resulte del “Programa” está protegido sólo si su contenido constituye una “obra basada en el Programa” (independientemente de haber sido creado por el “Programa” que corre.) El que esto ocurra de esa manera depende de lo que el “Programa” haga.

1. Está permitido copiar y distribuir por cualquier medio copias fieles del código fuente del “Programa” tal y como fue recibido, siempre y cuando se publique en

cada copia, de manera conspicua y apropiada, el aviso apropiado de derechos de autor y la renuncia a responder por la garantía correspondiente al “Programa”, se mantengan intactos los avisos referentes a esta licencia y a la respectiva ausencia de cualquier garantía; y se entregue a los receptores del “Programa” una copia de esta licencia.

Exigir una remuneración por el acto físico de transferir una copia está permitido; asimismo, también está permitido ofrecer una garantía a cambio de una contraprestación.

2. Está permitido modificar la copia o copias del “Programa” o cualquier parte del mismo, creando de esta forma, una “Obra basada en el Programa.” Asimismo, está permitido copiar y distribuir las modificaciones antes mencionadas o la obra misma bajo los términos de la Sección 1 mencionada anteriormente, y siempre y cuando se cumplan de igual manera las condiciones siguientes:

·a) Colocación de avisos, en la obra misma y por parte de quien realiza las modificaciones, en los que se informe que los archivos fueron modificados y la fecha de esas modificaciones.

·b) Otorgamiento de una licencia bajo los términos establecidos en esta Licencia Pública General que abarque la obra en su totalidad y sin cargo a terceras personas para el caso en el que se distribuya o publique una obra que contenga todo o parte del “Programa” o que constituya una obra derivada del mismo.

·c) Si el programa modificado normalmente lee comandos de manera interactiva cuando corre, cuando empieza a correr con dicho propósito interactivo, es necesario que aparezca un aviso que incluya la leyenda de derechos de autor correspondiente, así como la ausencia de responsabilidad por la garantía. Asimismo, dicho aviso deberá establecer que los usuarios de dicho programa tienen autorización para redistribuirlo bajo las mismas condiciones en las que les fue distribuido y les deberá informar cómo podrán tener acceso a una copia de esta licencia. (La excepción a esta condición tiene lugar cuando se trata de una “Obra basada en un Programa” que es en sí mismo interactivo, pero no envía normalmente un aviso.)

Las condiciones antes mencionadas se aplican a las obras modificadas como un todo. En el caso en el que las secciones de dicha obra que no se deriven del “Programa” sean identificables y razonablemente independientes y puedan separarse entre ellas, esta licencia y sus términos no se aplicarán a dichas secciones cuando éstas sean distribuidas como obras separadas. Sin embargo, cuando esas

mismas secciones se distribuyan como parte de la “Obra basada en el Programa”, dicha distribución deberá hacerse de acuerdo a los términos de esta licencia, cuyas autorizaciones para otros licenciarios tendrán los mismos alcances, sin importar qué parte creó quién.

Por medio de esta sección no se pretende exigir derechos o impugnar los derechos originados de una obra creada en su totalidad por otra persona, sino más bien se tiene como finalidad ejercer el derecho de controlar la distribución de obras derivadas o colectivas basadas en el “Programa”.

Asimismo, la sola inclusión de otra obra que no se base en el “Programa” aunada al “Programa” (o a una “Obra basada en el Programa”) dentro de un medio de almacenamiento o distribución no provoca que dicha obra deba registrarse por esta licencia.

3. Copiar y distribuir el “Programa” (o una “Obra basada en el Programa” de acuerdo a la sección 2), bajo los términos de las secciones 1 y 2 mencionadas anteriormente, ya sea en código objeto o en su forma ejecutable está permitido, siempre y cuando dicho “Programa” se acompañe también por cualquiera de los siguientes:

·a) El código fuente respectivo completo y legible por una máquina, el cual debe ser distribuido bajo los términos establecidos en las secciones 1 y 2 mencionadas anteriormente y a través de un medio normalmente usado para el intercambio de software;

·b) Una oferta por escrito y con una validez mínima de tres años, de proporcionar a cualquier tercera persona, por una cuota que no exceda el costo del acto físico de distribuir, bajo los términos de las secciones 1 y 2 antes mencionadas; y a través de un medio normalmente usado para el intercambio de software; una copia del respectivo código fuente completo y legible por una máquina; o,

·c) Toda la información recibida respecto a la oferta de distribución del código fuente correspondiente. (Esta alternativa está permitida únicamente para distribuciones no comerciales y siempre y cuando el “Programa” se haya recibido en código objeto o en forma ejecutable junto con esta oferta de acuerdo a la subsección b antes mencionada.)

El código fuente de una obra se refiere a la forma preferida para hacerle modificaciones. En una obra ejecutable, el código fuente completo se refiere a todo el código fuente de todos los módulos que contiene, además de cualquier archivo

de definición de interfaz asociado y de los **scripts** utilizados para controlar la compilación e instalación del ejecutable. Sin embargo, como una excepción especial, el código fuente distribuido no debe incluir cualquier cosa que sea normalmente distribuida (ya sea en forma de binarios o de código fuente) con los principales componentes del sistema operativo (como compilador, kernel, etc.) sobre el cual el ejecutable corre, a menos que el mismo componente acompañe al ejecutable.

Si la distribución del ejecutable o del código objeto se lleva a cabo mediante el ofrecimiento de acceso a una copia en un lugar designado, el ofrecimiento de acceso al código fuente en el mismo lugar equivale a la distribución de dicho código fuente, aun cuando terceras personas no estén obligadas a copiar el código fuente junto con el código objeto.

4.El “Programa” no puede copiarse, modificarse, sublicenciarse ni distribuirse a menos que se haga bajo los términos y condiciones de esta licencia. Cualquier intento por hacer lo anterior de otra forma, será nulo y extinguirá automáticamente los derechos surgidos de esta licencia. Sin embargo, las licencias de las personas que hayan recibido copias o derechos bajo esta licencia, seguirán vigentes mientras dichas personas cumplan con sus obligaciones.

5.Mientras no se firme la presente licencia no existe obligación de aceptarla. Sin embargo, no existe autorización, y por lo tanto está legalmente prohibido, modificar o distribuir el “Programa” o una “Obra basada en el Programa” a menos que se acepten los términos y condiciones de la presente licencia. Por lo anterior, del acto de modificar o distribuir el “Programa” o una “Obra basada en el Programa” se presume la aceptación de los términos y condiciones de la presente licencia para copiar, distribuir o modificar dicho “Programa” u “Obra basada en el Programa”.

6.Cada vez que se distribuya el “Programa” (o cualquier “Obra basada en el Programa”), quien recibe la copia del mismo recibe también, de manera automática una licencia de parte del licenciante original para copiar, distribuir o modificar el “Programa” bajo los términos y condiciones de esta licencia. No podrán imponerse más restricciones al ejercicio de los derechos del licenciatario que los establecidos en esta licencia. Quien distribuye el “Programa” no es responsable por el cumplimiento de la presente licencia por parte de terceras personas.

7.En el caso en el que como consecuencia de orden judicial o de las pretensiones demandadas por violación a una patente o por cualquier otra razón (de manera enunciativa, mas no limitativa) se imponen condiciones (ya sea por orden judicial,

contrato o por otro medio) que se contradicen con las condiciones de esta licencia, estas últimas no se eximen de su cumplimiento. Como consecuencia de la imposibilidad de cumplir con ambas obligaciones mencionadas, el “Programa” no podrá distribuirse. Por ejemplo, si una licencia de una patente prohíbe la redistribución gratuita del “Programa” por parte de quienes reciben copias del mismo de manera directa o indirecta, entonces la única forma de cumplir con ambas licencias, ésta y la de la patente, será abstenerse de distribuir el “Programa”.

En el caso en el que cualquier parte de esta sección sea declarada inválida o inexigible bajo cualquier circunstancia particular, el resto de la misma continuará surtiendo sus efectos para esa circunstancia, al igual que la sección en su totalidad para las demás circunstancias.

El propósito de esta sección no es inducir a la violación de patentes o del ejercicio de otros derechos intelectuales, como tampoco impugnar la validez de tales demandas por incumplimiento, sino mas bien, pretende proteger la integridad del sistema de distribución del Software Libre, el cual consiste en la práctica y uso de licencias públicas. Mucha gente ha hecho generosas contribuciones a paquetes de software distribuidos bajo este sistema confiando en la aplicación de dicho sistema; y es decisión del autor/donante distribuir el software a través de cualquier otro sistema sin que un licenciatarario pueda interferir en esa decisión.

Esta sección pretende aclarar todo aquello que se considera consecuencia del resto de esta licencia.

8.En el caso en el que la distribución y/o uso del “Programa” esté restringida en ciertos países, ya sea por patentes o interfaces protegidas por el sistema de propiedad intelectual, el titular original de los derechos de autor del “Programa” que lo sujeta a esta licencia tiene la facultad de agregar una limitación de tipo geográfico a la distribución, por virtud de la cual se excluya a dichos países; de manera que la distribución del mismo se permita únicamente en los países no excluidos. En este caso, dicha limitación se tiene como parte integrante de esta licencia.

9.Es facultad de la Free Software Foundation publicar, en cualquier momento, tanto versiones revisadas como versiones de reciente creación, de la Licencia Pública General. Las versiones nuevas pueden diferir en detalles a fin de afrontar y resolver nuevos problemas o preocupaciones, pero conservando siempre el espíritu de la presente versión.

Cada versión tendrá asignado un número. En el caso en el que el “Programa”

especifique un número de versión de esta licencia para su aplicación y además, incluya la frase “y cualquier versión posterior”, el licenciatarario podrá sujetarse, a su elección, a los términos y condiciones de la versión expresamente mencionada o de cualquiera de las versiones posteriores de la misma publicadas por la Free Software Foundation. Por otro lado, en el caso en el que el “programa” no especifique un número de versión de licencia, el licenciatarario podrá elegir cualquier versión que haya sido publicada por la Free Software Foundation.

10. En el caso en el que se deseen incorporar partes del “Programa” a otros paquetes de Software Libre cuyas condiciones de distribución difieran a estas, es necesario solicitar permiso por escrito al autor. Cuando se trate de software cuyo titular de los derechos de autor correspondientes sea la Free Software Foundation, la solicitud de permiso deberá dirigirse a ésta última, quien en algunas ocasiones hace excepciones como esta. La decisión emitida por la Free Software Foundation se basará tomando en cuenta la finalidad de preservar el estatus libre de todos los derivados del Software Libre y de promocionar que se comparta y se reutilice el software en general.

EXCLUSIÓN DE GARANTÍA

11. COMO CONSECUENCIA DE QUE EL “PROGRAMA” SE LICENCIE COMO GRATUITO, EN LA MEDIDA EN QUE LA LEY APLICABLE LO PERMITA, NO EXISTIRÁ GARANTÍA ALGUNA POR LA QUE SE DEBA RESPONDER. SALVO DISPOSICIÓN ESCRITA EN CONTRARIO, LOS TITULARES DE LOS DERECHOS DE AUTOR RESPECTIVOS Y/U OTRAS PARTES PONEN A DISPOSICIÓN EL “PROGRAMA” SIN GARANTÍA DE NINGÚN TIPO, EXPRESA O IMPLÍCITA, INCLUYENDO DE MANERA ENUNCIATIVA MAS NO LIMITATIVA, LAS GARANTÍAS IMPLÍCITAS DE TIPO COMERCIAL U OTRAS INHERENTES A ALGÚN PROPÓSITO ESPECÍFICO. EL RIESGO DE QUE EL “PROGRAMA” ESTÉ EN PERFECTAS CONDICIONES Y FUNCIONE TAL Y COMO DEBE FUNCIONAR CORRE POR CUENTA DE QUIEN LO RECIBE, AL IGUAL QUE LOS GASTOS NECESARIOS PARA SU SERVICIO, REPARACIÓN O CORRECCIÓN EN EL DADO CASO EN EL QUE DICHO “PROGRAMA” CONTENGA DEFECTOS.

12. A MENOS QUE ASÍ LO DISPONGA LA LEY APLICABLE O EXISTA ACUERDO ESCRITO EN CONTRARIO, NINGÚN TITULAR DE LOS DERECHOS DE AUTOR O PERSONA FACULTADA, SEGÚN LAS SECCIONES ANTERIORES DE LA PRESENTE, PARA MODIFICAR Y/O DISTRIBUIR EL “PROGRAMA” SERÁ RESPONSABLE POR LOS DAÑOS

YA SEAN GENERALES, ESPECIALES, INCIDENTALES O CONSECUENCIALES RESULTADO DEL USO O INCAPACIDAD DE USO DEL “PROGRAMA” (INCLUYENDO DE MANERA ENUNCIATIVA MAS NO LIMITATIVA LA PÉRDIDA DE INFORMACIÓN, INEXACTITUD EN LA INFORMACIÓN, PÉRDIDAS SUFRIDAS POR EL USUARIO DEL “PROGRAMA” O POR TERCERAS PERSONAS O LA INCAPACIDAD DEL “PROGRAMA” PARA OPERAR CON OTROS PROGRAMAS), AUN CUANDO DICHO TITULAR O CUALQUIER OTRA PERSONA HAYA ADVERTIDO DICHA POSIBILIDAD DE DAÑO.

FIN DE LOS TÉRMINOS Y CONDICIONES

Cómo aplicar estos términos a los nuevos “programas”

En el caso en el que se esté desarrollando un “Programa” nuevo y se tenga la intención de hacerlo de uso público, la mejor forma de lograrlo es haciéndolo Libre, y de esta forma, permitir a cualquiera redistribuirlo y cambiarlo bajo los términos y condiciones de esta Licencia.

A fin de lograr lo anterior, se deben incluir los siguientes avisos al “Programa”. Es más seguro incluir dichos avisos al principio de cada archivo de código fuente para aclarar de manera más eficiente la exclusión de garantía mencionada. Asimismo, cada archivo debe tener por lo menos la frase “Derechos Reservados” (Copyright para los países con ese sistema) relativa a los derechos de autor, así como la referencia al lugar donde se encuentre la leyenda y especificaciones completas de los mismos.

La referencia al nombre del “Programa” y a una idea de lo que el mismo hace.

Derechos Reservados © yyyy nombre del autor

Este es un Software Libre; como tal redistribuirlo y/o modificarlo está permitido, siempre y cuando se haga bajo los términos y condiciones de la Licencia Pública General GNU publicada por la Free Software Foundation, ya sea en su versión 2 ó cualquier otra de las posteriores a la misma.

Este “Programa” se distribuye con la intención de que sea útil, sin embargo carece de garantía, ni siquiera tiene la garantía implícita de tipo comercial o inherente al propósito del mismo “Programa”. Ver la Licencia Pública General GNU para más detalles.

Se debe haber recibido una copia de la Licencia Pública General GNU con este “Programa”, si este no fue el caso, favor de escribir a la Free Software Foundation,

Inc., 59 Temple Place – Suite 330, Boston, MA 02111-1307, USA.

Asimismo, se deben incluir las direcciones de correo electrónico y convencional del autor del “Programa” a fin de contactarlo.

En el caso en el que el “Programa” sea interactivo se debe incluir un aviso cuando inicie el modo interactivo como el siguiente:

Gnomovision versión 69, Derechos Reservados © primer año de publicación y nombre del autor

Gnomovision carece totalmente de garantía; para más detalles teclee “show w”. Este es Software Libre y está permitido redistribuirlo bajo ciertas condiciones; teclee “show c” para más detalles.

Los comandos hipotéticos “show w” y “show c” deberán desplegar las partes correspondientes de la Licencia Pública General. Obviamente, los comandos que se utilicen pueden ser distintos a los mencionados, inclusive pueden ser clicks del ratón o elementos del menú, según sea más conveniente.

Asimismo, el autor del “Programa” debe obtener de su empleador (en caso en que dicho autor trabaje como programador) o de su escuela, si ese es el caso, una renuncia firmada a los derechos de autor por el “Programa”, si es que fuera necesario. El siguiente es un ejemplo:

Yoyodyne, Inc., por medio de la presente se renuncia a cualquier derecho de autor que corresponda al programa “Gnomovision” cuyo autor es James Hacker.

Firma Ty Coon, Presidente de Vice, 1 de Abril de 1989.

Esta Licencia Pública General prohíbe incorporar el “Programa” a programas propietarios. En el caso en el que se trate de un programa que a su vez sea una librería de subrutinas, es más conveniente permitir ligas de aplicaciones propietarias con la librería. Si es esto lo que se desea, entonces se debe usar la Licencia Pública General Menor de GNU, en lugar de esta licencia.

Anexo 5

Subsistema de globos piloto.

Introducción

Por sus características este subsistema se dividió en dos módulos, uno para la captura de datos a ser ejecutado conectado al teodolito y para correr en ambiente Microsoft Windows, y otro para la consulta de datos, vía WEB, en el nuevo servidor de CENICA.

Compilación y ordenamiento de las bases de datos existentes.

El CENICA entregó a Matías Software Group un CD con datos agrupados en tres directorios: C1 (Campaña del 2000), C2 (Campaña del 2001), C3 (Campaña del 2002), KyToon01 (Pruebas con este tipo de globo) y MitcamDF.

Los datos entregados venían en diversos formatos derivados del utilizado por el teodolito TD3 para sus comunicaciones. Las únicas observaciones que dieron problemas irre recuperables fueron C1/01021213.AZC y C2/00071213.UNI, ya que el teodolito sufrió alguna falla y comenzó a reportar números incoherentes. Recomendamos poner atención a futuras fallas del teodolito, ya que puede tener algún error en la programación de la electrónica, aunque por el bajo grado de incidencia, parece ser un problema que aparecerá en muy pocas ocasiones en la práctica.

Ya normalizados los datos en un formato uniforme, se subieron a la base de datos del SINAICA, en la tabla *globo.obs*, con la siguiente estructura:

Columna	Tipo	Modificadores
fecha	timestamp without time zone	
lugar	text	
comm	text	
interv	integer	not null
nmuest	integer	not null
func	text	
azimuth	real[]	not null
elev	real[]	not null
visible	boolean[]	
discz	real[]	

Restricciones de prueba:

```

"$1" CHECK (array_upper(elev, 1) >= nmuest AND array_upper(azimuth,
1) >= nmuest)
"obs_azimuth" CHECK ((0::double precision <= ALL (azimuth)) AND
(360::double precision >= ALL (azimuth)))
"obs_elev" CHECK ((0::double precision <= ALL (elev)) AND
(360::double precision >= ALL (elev)))
"obs_interv" CHECK (interv > 0)
"obs_nmuest" CHECK (nmuest >= 2)

```

De especial importancia es mencionar que la campaña a la que pertenece una observación dada está indicada en el campo *comm*.

Interfase teodolito-base de datos

Se acordó realizar una interfase que funcionara bajo la plataforma Win32, para reducir los requerimientos de las portátiles que corrieran el programa de captura. El programa fue realizado en Perl y viene dentro de un ejecutable de instalación para Windows en el disco compacto de entrega, junto con documentación, código fuente y especificaciones.

Para lograr la comunicación con el teodolito, se realizó una investigación partiendo de la documentación de la llamada carpeta y un aparato para desarrollo de prototipos seriales, resultando un cable intermedio exprofeso para conectar al teodolito con la computadora, usando el cable teodolito-carpeta como único intermediario.

Generación de vectores de viento.

Tras un pequeño análisis trigonométrico, se llegó a las siguientes conclusiones:

Sean A el ángulo azimutal, E el ángulo de elevación registrados por el teodolito, T el instante en el tiempo asociado por el registro en cuestión y z(t) la función de altura respecto al tiempo estimada,

si H es la magnitud de la proyección del vector (A, E, z(t)) al plano z=0, entonces $H = z(T) / \tan E$

y por lo tanto, finalmente, $X = H (\sin A)$, $Y = -H (\cos A)$.

Aunque falta el dato de la posición inicial del globo, esta en principio es irrelevante, especialmente debido a que lo que se busca es la derivada, o sea, los deltas, de las posiciones del globo, donde la constante inicial se pierde. Por simple

cálculo diferencial, se deduce que los valores de los vectores resultantes de las deltas de las posiciones del globo se mantienen, aun cuando la constante inicial tienda a cero, y por lo tanto queda fuera del problema.

Esta solución quedó registrada como una función en la base de datos, llamada `globo.xyz`, cuya declaración queda definida en el archivo `SINAICA_def.sql` del disco compacto de entregas. También, para mayor comodidad, se incluyó una vista `globo.winds` que entrega los vectores de viento ya calculados, disponibles a través de SQL.

Aplicación de análisis y reporte.

Se implementó una página Web que presenta las observaciones disponibles con sus datos, y un generador de gráficas SVG por Web de trayectorias y campos de viento de las observaciones.

