Daily System Alerts Development Information
SIAD for the IDEAM

Samir Ernesto Barros Palencia, Camilo Andrés Palacios Cubillos
System engineering Program, Engineering Faculty, Universidad Católica de Colombia
Bogotá D.C., Colombia

Abstract: This article presents the information process development for data administration which appears on the daily report of the IDEAM, this document notifies to each one of risk entities about any possible alert that can come around in the township and that is an excellent advanced because it has been done hand-made and inadequate way. In the same research is showed a Data Warehouse design which main purpose is generate historical reports giving in this way the opportunity for generating reports about possible standards of meteorological behavior in the country and with them try to help in making decisions for avoiding natural disasters.

Keywords: RUP (Rational Unified Process), IDEAM, Data Warehouse, Information System

I. INTRODUCTION

Forecast and alerts service offers by “Instituto de Hidrología, Meteorología y Estudios Ambientales” (IDEAM) has become the most important offered by this entity because it generates a report call “daily alert report” and this have an important function: provide information to different governmental entities (Attention and prevention of disasters and civil defense) about any hydrometeorology event which can appear in different place in the country.

Also, this report is revised by each entity in charged around the country and that is a very useful tool when making convenient and true decisions about the steps to follow and in that way face any risk or alert (red, orange or yellow) that can have place in any region.

Taking into account the previous ideas, is very important for the forecast and alerts office (OSPA) in the IDEAM, which is the entity in charge of this document and nowadays is in a certification and accreditation process, have a platform where all the information about weather, hydrology and environmental supplied by any region in the country can be managed. That information is collected in a database call “Subsistema de hidrología y meteorología” (SSHM).

With the aim of optimizing information management which contains the daily alert system it was designed and implemented an information system that is supported on an online transaction process database scheme (OLTP), which is fed from SSHM, and in that way can generate accurate, true, easy and understandable reports.

Besides, it is a possibility for generating historical reports in order to detect any environmental behavior in the country and in this way trying to avoid a catastrophe. For that, it was designed with the information system a data warehouse (OLAP database) and it has the purpose of establish solid bases for its future implementation.

In this article, is possible to know all the steps did in this research project, which is divided into three sections. The first part is about information system and Data Warehouse with the aim of showing its organization, function and importance for any enterprise.

In the second part, is possible to observe how the information system was done, it is also presented the SIAD development where all the requirements and design were done in order to make the implementation and testing phases.
In the third part, is showed the Data Warehouse design step by step and also how to reach to the multidimensional model.

II. THEORETICAL FRAMEWORK

a. Information system

Taking into account the concept for the FIB (Facultad de Informática de Barcelona de la Universidad Politécnica de Cataluña), information systems use information technologies to provide people with the information they need, giving vital support to business operations, management and decision-making. Enterprises use information systems as strategic tools for innovating, competing and achieving their goals in a globalized market. Information systems integrate people, processes, data and technology and extend beyond traditional organizational boundaries to establish more efficient relationships with suppliers, distributors and customers [1].

In the same way, for the Münster University Information System Department, the information systems are focus in the applying of information and communication technologies in order to support business possibilities and process [2].

b. Warehouse Data

In this topic is necessary have in mind the computer scientist Bill Inmon and the expert Ralph Kimbal. For the first one, a data warehouse is a data collection classified by topics, integrate, time-variant and nonvolatile which helps in decision making process for people who manage an enterprise.

It is also important say that for Imon, considered the Data Warehouse’s father, the data of a warehouse are integrated because all of them come from different sources inside an organization. Additional, it is important to know that data are nonvolatile, it means they are not actualized in real time but it is done periodically depending on the operational systems. Taking into account data are classified by topics, Inmon says that a warehouse is organized from the most to the less important data and for concluding he said that data are variables in time because they are changeable through time and they are precised and valid only for a short period of time [3].

For the second one, the expert in Data Warehouse, Ralph Kimball, the data in a warehouse is a copy from transactional and structured information only for looking up and analyzing data.

According to Kimball, best known as Data Warehouse’s original architect, the original unit in that place is a data mart, which definition is an information repository pointed to a specific area or department in any organization. For him, in a Data Warehouse there are four main ideas: focus on business’ needs, build a proper infrastructure, deliver in a growth way and offer a complete answer [4].

III. SIAD’S DEVELOPMENT PHASES

Product development cycle was guided by the RUP (Rational Unified Process) methodology and it also applied the UML (Unified Modeling Language) like a representative language for all the artifacts developed in this project. That methodology was chosen because the IDEAM has a software which works with that.

The RUP methodology is divided into four phases (inception, elaboration, construction and transition) in each one of them some iterations can be done, and can perform some disciplines like business modeling, requirements, analysis and design, implementation, test and deployment, according to the milestone which has been marked for each iteration.
In this step system functionalities were agreement, it was after some meetings with the stakeholders and it was necessary different versions for getting the approval for the IDEAM.

For requirements document it was applied a template done by the IDEAM and it is showed in table I.

<table>
<thead>
<tr>
<th>IDENTIFICADOR REQUERIMIENTO:</th>
<th>NOMBRE:</th>
<th>&lt;&lt;Nombre del requerimiento funcional&gt;&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLASIFICACIÓN:</td>
<td></td>
<td>&lt;&lt;El nombre de cada opción, componente o proceso dentro del SIAD, ejemplo consulta, reporte, etc.&gt;&gt;</td>
</tr>
<tr>
<td>PROPÓSITO:</td>
<td></td>
<td>&lt;&lt;Objetivo general del requerimiento&gt;&gt;</td>
</tr>
<tr>
<td>DESCRIPCIÓN:</td>
<td></td>
<td>&lt;&lt;Descripción detallada del requerimiento&gt;&gt;</td>
</tr>
<tr>
<td>CONSIDERACIONES:</td>
<td></td>
<td>&lt;&lt;Descripción de los aspectos del requerimiento, teniendo en cuenta el alcance del proyecto y detallando las características no contempladas en el mismo&gt;&gt;</td>
</tr>
<tr>
<td>NIVEL CVS</td>
<td></td>
<td>&lt;&lt;Especificar el nivel de CVS del requerimiento funcional. Puede ser: aplicación, módulo o submódulo.&gt;&gt;</td>
</tr>
</tbody>
</table>

3)  **System design stage**

In the first meetings done at the IDEAM more than the functional requirements were identified, it was also found a use cases list that any user can have inside the system and thanks to that it was possible to had the architecture baseline for the project.

b. **Elaboration phase**

In this phase all the user requirements were analyzed and the system architecture baseline was defined, for that, it was taken all the previous information about use case. At the end of this stage all the user requirements were ready for the construction phase.
In this phase two iterations were done. In the first one the requirement document was finished, including all the new requirements according to the system needs, also it was done the user interface and the quality attributes were taking into account for presenting the application.

The second one was the last modification for each use cases and the first version of the entity-relationship model, class, deployment and component diagrams was designed and it was essential for the next step that began with the approval for the IDEAM.

1) **Raising requirement step**

It was won thanks to the meetings at IDEAM’s offices where all the quality attributes were defined. This step was done using the format designed by the IDEAM as it is observed in table II.

<table>
<thead>
<tr>
<th>IDENTIFICADOR</th>
<th>NOMBRE:</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNF_##</td>
<td>&lt;&lt;Nombre del requerimiento no funcional&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLASIFICACIÓN:</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;Nombre del sistema al que pertenece el requerimiento no funcional, en este caso SIAD&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DESCRIPCIÓN:</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;Descripción detallada del requerimiento&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CONSIDERACIONES:</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;Descripción de los aspectos del requerimiento, teniendo en cuenta el alcance del proyecto y detallando las características no contempladas en el mismo&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NIVEL CVS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;Especificar el nivel de CVS del requerimiento no funcional. Puede ser: aplicación, módulo o submódulo.&gt;&gt;</td>
</tr>
</tbody>
</table>

2) **System design stage**

At the beginning of the stage, was done the use cases document using the template given by the IDEAM where it is described how the system user must work on it.

In the second iteration the system architecture document was done and for this the template given by the IDEAM was applied and it shows the information system structure in order to approve the development environment which includes the following diagram:

- System component diagram where all the logical distribution is showed.
- Deployment diagram where the physical or hardware distribution is showed.
- Class diagram where all system objects are described and the relationships among them.
- Entity-relationship model where all the entities and relationships needed for the system were showed.
- Data dictionary, where all the attributes of the database were showed, for this the format done by the IDEAM was applied. (see table III.).

<table>
<thead>
<tr>
<th>Table III. Data dictionary template</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMBRE OBJETO: &lt;&lt;nombre del objeto a crear&gt;&gt;</td>
</tr>
<tr>
<td>DESCRIPCION: &lt;&lt;descripción del propósito del objeto&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Columnas</th>
</tr>
</thead>
<tbody>
<tr>
<td>PK</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Relaciones:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Columnas</td>
</tr>
</tbody>
</table>

3) **Implementation stage**

With the aim of having a prototype of user interface all the needed web pages for the users requirements were applied and in that way it explained what any user must do. This prototype was developed and worked in order to describe each requirement and use case.

c. **Construction phase**
In this step all the previous work were finished and started the implementation of the software coding each use case spoken at IDEAM.

Here four iterations were done and each one of them was an operative version for the system, it means each iteration allowed to create a new and better version with more functions because RUP had its development in iterative and incremental versions.

1) Implementation stage

In this stage all the use cases were implemented and a software version that fulfill all the functional requirements was development.

The process began with the creation of the tables in the IDEAM database using Oracle 11g, and then the web application codification was done in the IDE (integrated development environment) Jboss Developer Studio 7.0.0 using JAVA as programming language.

The coding system began configuring the development environment. In the first moment it was used the IDE given IDEAM and it had some inconvenient with the JPA (Java Persistence API) version which comes pre-installed and it was necessary to update the Jboss Developer program from 6.0.1 version to 7.0.0 version and in that way is possible to select the JPA 2.1 version. There were still problems because it was necessary to write manually in the persistent file (persistence.xml) because the program did not match with JBOSS EAP-6 server.

After, the libraries for JSF (Java Server Faces) version 2.1 and IceFaces 3.3.0 were discharged and integrated. The versions were selected because are the latest and its compatible with the dynamic projects 3.0 and Mojarra 2.1.21.

When the development environment was configured each use case was codify in the phase iterations according to the priority and complexity given by the IDEAM.

In the first iteration use cases related to hydrology and meteorology were implemented. The first step was codify use cases CU_09 (write meteorological information) and CU_010 (write hydrology information). The codes correspond to meteorological and hydrological information which is going to be useful in daily alert report. Besides, the use case CU_015 was added (Modify threshold), that allows the modification of maximum and minimum temperature and precipitation volume.

The second iteration use cases about fires and landslides CU_011 (write fire information) CU_012 (write landslides information), those use cases correspond to the inclusion of information in the same database and is going to be useful for generating daily alert report.

In the same way some other cases were codified CU_06 (query by date range) and CU_07 (query daily). The first one is related to query maximum and minimum temperature and precipitation by date range parameter (threshold). The second one is related to static queries which are found in the report in order to alert experts about any change in the weather or any dangerous situation. The last use case in this iteration is CU_08 (load the document with the appropriate volume for the reservoirs) and this allows user to load an Excel file the proper volume for the reservoir which is necessary information in order to generate the daily alerts report

The third iteration worked with management module, the use case added were CU_014 (check the last access), CU_016 (add users to the system) and CU_017 (add a role to the user). For CU_014 a chart where all the actions that have been made in the system with its respective actor and date are going to be shown.

2) Test step
In this stage the functional tests system were performed and in this case unit test in order to check the correct operation system. In the applying of each use case it was tested and if there was a mistake the corresponding correction was made.

d. Transition phase

It began with the product delivery and the unit testing to IDEAM because they must verify its correct operation as final users. This phase had two iterations, the first one functional testing are taken and in the second one the program is put into production.

1) test step

In this step the functional testing by IDEAM will be made in order to verify the unit tests done by engineers. During these tests all changes and functional improvements determined by the IDEAM will be do in order to be sure about the proper software operation.

After this, the final test will be execute with the final user in order to correct possible mistakes and also get the approval by Alberto Pardo the alert coordinator IDEAM.

The functional test and the final user test will be do at office computer IDEAM, it is expected to finish on June and in that way begin with the deployment step.

2) Deployment step

When the software will be approved by the engineer Alberto Pardo this stage is going to have place and this is going to be done by the computer office which is responsible for the management of IDEAM information systems.

This stage will start when all the tests are approved, so it is not possible to give an exact date for starting the software because it is attached to IDEAM’s time.

IV. DATA WAREHOUSE DESIGN

The relational databases have been growing in popularity in recent decades; they are so important that most of the enterprises and companies store their data in these systems. Database management systems (DBMS) have been used to support online transaction processing (OLTP) because companies do a lot of transactions CRUD (create, read, update and delete) for their business daily.

Furthermore, Data Warehouses feeds from the OLTP database with historical information becomes a valuable resource that supports the information system for decision-making and enables extracting knowledge from such historical information to analyze, formulate and implement the strategies and achieve organizations’ objectives.

Since the acquisition by the IDEAM from the relational database to the present time it has been gathering an increasing amount of data, which are mostly hydrometeorology indicators which report different weather behaviors in the country.

Taking into account the previous information, it was necessary to design an application that puts into practice those historical operational data and help in decision making, as a method to detect patterns in weather behavior. With that aim a Data Warehouse (multidimensional database) was designed where all the indicators that are used in daily alerts can be stored and for getting this Kimball’s methodology is going to be used.

a. Design methodology

Kimball’s methodology based its process in four steps which together make the Data Warehouse design becomes a globally consistent warehouse [5], in addition, it is recommended the maximum duration for the data warehouse. The steps for the data warehouse design are:

- **Process Selection:** The process refers to the objective in a particular data item market [3].
- **Granularity Selecting (detail level):** It means saying what is going to represent each record in the fact table [3].
• **Dimensions’ identification and creation:** The dimensions set the context for asking on the fact table. A well-constructed set of dimensions makes the store understandable and easy to use [3].

• **Events Selection:** The granularity in the fact table indicates that measures can be used; the attributes that the fact chart collects are called measures. All facts should be expressed in the chosen granularity level [3].

  o **Pre-calculated values storage in the fact table:** After selecting the facts, it is necessary to re-examine them in order to determine if there is a possibility for using pre-calculated values [3].

The last part is going to be the **database duration** which measures until the past time that the fact chart must be last. Storing historical information can range from 2 to 10 years depending on the business [3].

**b. Methodology applying design**

1) **Process selection**

The purpose in this data warehouse project is to take weather conditions in order to get information from all the historical information that is stored daily in the IDEAM.

This process was seen as the most important because the current weather change, and is necessary for the country to have a data warehouse and in this way analyze weather behavior in Colombia in order to make accurate and timely decisions that benefit the entire population.

2) **Granularity selecting**

Each record in the fact chart will represent the meteorological behavior (maximum temperature, minimum temperature and / or precipitation) in the country.

3) **Dimensions’ identification and creation**

Dimensions were identified thanks to the granularity level. First, the need to restrict the queries on the fact chart by **Lugar**, **Tiempo**, **Estación** and **Zona**.

**Lugar** dimension was established in order to restrict the query according to the place where the measure was made either by region, department or municipality. Similarly dimension **Zona** indicate a point but in this case separating them by rivers or islands.

On the other hand, there is the **Estación** dimension that will be allowed to restrict queries by name or height, as the weather behavior is different at each height. And finally the **Tiempo** dimension which will see indicators in different time ranges given.

To conclude the dimension chart was added in order to provide a perspective or analysis on the form of a measurement made on a table. After this step the dimensions were in this way:

<table>
<thead>
<tr>
<th>Table IV. Data Warehouse Dimensions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dimension</strong></td>
</tr>
<tr>
<td>Lugar</td>
</tr>
<tr>
<td>Tiempo</td>
</tr>
<tr>
<td>Estación</td>
</tr>
<tr>
<td>Zona</td>
</tr>
</tbody>
</table>

4) **Event selecting**

For the event selecting it was taking into account the granularity and the chosen process so measures were taken (attributes) from the fact chart: Maximum temperature, minimum temperature and precipitation volume.

With these facts and historical information stored in the fact chart they will be available to check weather behavior that has been in the country in recent years.

5) **Database length**

As climate change is weather modification over the years, it is necessary get useful knowledge from the data store and save information from
meteorological indicators from 8 to 10 years because if someone wants to know any weather behavior is necessary to compare weather from several years.

c. **Extraction, transformation and load process (ETL)**

The ETL process involves the migration process from OLTP database to the Data Warehouse. Extraction is the first step where the information can be search in the transactional database. Then comes the transformation phase where rules are applied to data in order to proceed to the final stage of the process, this is to load the data in the warehouse [6].

1) **Extraction process**

The extraction process is the first phase of data preparation, it details the corporate database (OLTP) which are thought to feed the data warehouse. These extracted data is stored in a repository or view, in which the data will undertake all the necessary processes to load the warehouse [3].

To start it is going to be explain where information comes from to fill the dimensions of the data warehouse from the OLTP database.

- Data information entry for the Estación dimension is going to be done on the SIOV_Estaciones chart in the OLTP database where the Altitud and nombre_es attributes information are going to be taken.
- For the Zona dimension data is going to be taken from the information in the chart SIOV_Estaciones, of the attributes nom_zona and nom_subzona.
- For data dimension Lugar is going to be taken from nom_area of the SIOV_Estaciones table and desc_departamento and desc_municipio from SIOV_Divipolas table.
- The Tiempo dimention data is going to be taken from CSMT_fecha from SHMT_Cuasirreal_meteorologia table.
- The last one, fact dimension is going to be taken from CSMT_tmp_max, CSMT_tmp_min and CSMT_prec.

2) **Transformation process**

In the transformation stage a set of business rules or functions are applied to the extracted data in the previous phase, among these rules are, standardize measurements, dates, inconsistencies formats; avoid redundant data and treating null values, among others [6].

In the ETL process for IDEAM Data Warehouse only will be transform the data for Tiempo and Zona charts. For the Zona table the records “don’t apply” will be apply for the null information, this with the aim of optimizing data store.

On the other hand, Tiempo chart is going to be separated from data with the name CSMT_fecha for making easy to type them into the chart.

3) **Charge process**

In this phase, the data were extracted, cleaned and structured in the previous phases is loaded into the Data Warehouse. Depending on the requirements of the organization, this process may involve a variety of different actions [7].

For this process the data resulting from the extraction and modification made on the OLTP database in the data store was loaded. For getting this aim it is advisable to make the charge on weekends, as the IDEAM handles a large number of transactions this process during office hours can make the OLTP database work slowly and that delay the processes that depend on it.

d. **Multidimensional model**
V. CONCLUSIONS

At the inception phase great importance of business steps modeling and requirements were seen because it was recognized the importance of IDEAM, known as the responsible for checking weather and warnings in order to avoid disasters that can affect the population. In the same way, it was detected the poor technological infrastructure that government entities have in Colombia because important documents such as the daily alerts report are done manually.

On the other hand, in the requirement stage was evident that the requirements and use cases are very important because with them all the project was done and all needs are presented in detail in the software elaboration. Without them it would have been impossible to develop system information because any bad requirement or data can result in an unexpected product and it implies to re-organize the entire project again.

In the same way, it was showed the advantage of having a permanent contact with the client for the proper development of the project like is stated in RUP as any concerns about the company’s business are quickly solved. This helps and is useful because an excellent job is going to have place.

On the other hand, it was evident that follow a development methodology facilitates any project design as they use one that provides guidelines used by experts facilitates their implementation and ensures the delivery of a product that satisfies the customer.

However, it was evident that not having enough experience in the design and development of software can choose frameworks that are not compatible with the development environment, which makes the implementation phase delayed and also the project. But the use of frameworks facilitates the development and implementation of any software project, as long as some of these basics are taken, otherwise, the development team is delayed while the use of these learning. But at the end is useful because it ensures optimal quality and software.

It is also important to say that having implemented in an unknown environment cause a delayed in the early stages of the project, it gave because more time than expected was spent in how to use. But when the members of the team have different knowledge it was easier to develop the project because all the tasks were discussed and assigned and is better to have work team than an individually job.

Talking about Data Warehouse, it was concluded that by implementing storage systems for the management and organization of data a very useful tool is the Data Warehouse where information can be modeled in an orderly way which are part of the organization for gathering weather information.

In the same way, the application of this storage structure, facilitates, organizes and systematizes processes that the organization performed intuitively and unreliable manner, taking advance in making decisions affecting the Colombian population.

Finally, it should be noticed that the SIAD will impact positively on the interested areas at IDEAM, as it will facilitate the completion of the report, streamlining the process and allowing, if necessary, conducting more than one report per day. Also benefit the Colombian population, which thanks to this new information system will be able to see more accurate and timely reports.

ACKNOWLEDGEMENT
Our deepest and sincere thanks to all those who have collaborated with their help in this project.

REFERENCES


