



## Pre announcement of an open source data warehouse management system

In the past Nippur mentioned several times that development of data warehouses is a time consuming and expensive activity. This is caused by the fact that most work is done manually. Both the creation of robust data models as well as the tedious development of accompanying load processes. On a regular basis we pleaded for the use of so called data warehouse management systems. As a consultancy firm Nippur is used to working with products that are available for this purpose. We therefore understand the advantages these products offer.

What a pity that choice is limited. Solutions that exist are extremely expensive. As if they are only meant to be used by the happy few: large multi-national organizations with big BI competence centers.

Thanks to the invention of the [Data Vault](#) modeling concept as conceived by Mr. Dan Linstedt an interesting new situation arose. This concept is so generic that it can be automated. We know how true this is because we helped many customers in implementing a Data Vault oriented data warehouse. Each time we developed tools (script generators) to automate the repetitive (boring!) tasks we saw the advantages. One of which is a true acceleration in the implementation of a proper data warehouse architecture. In some cases up to 80% faster. With the same budget customers can investigate and implement more business cases. And if you ask us that is what it is all about. It is also more interesting and hence more fun to do. Understanding the customer, helping to define and implement relevant business cases.

This experience and the fact that existing data warehouse management systems are too expensive made us decide to develop such a system and distribute it based on an open source license model to everybody in need of a data integration platform. This project and resulting product is called Quipu. The name Quipu is derived from the Quipu's archeologists found in and around Cuzco (also written as QOSQO) in Peru.

Quipu's or khipus (sometimes called talking knots) were recording devices used in the Inca Empire and its predecessor societies in the Andean region. A quipu usually consisted of colored spun and plied thread or strings from llama or alpaca hair. It could also be made of cotton cords. The cords contained numeric and other values encoded by knots in a base ten positional system. Quipu's might have just a few or up to 2,000 cords. [Source: Wikipedia].



Figure 1. Example of a Quipu.

Soon version 1.0 of Quipu will be born. Hence this pre announcement.

### Quipu

In august 2008 QOSQO was founded by Nippur to provide on the one hand data warehouse services (based on data warehouses as developed by sister Nippur). And to develop and launch Quipu, the open source data warehouse management system on the other hand.

We already know about open source BI tools that offer functionality in the area of reporting, ETL, databases, data mining, wikis etcetera. However the data integration component is the missing component. Probably because of the specific knowledge required to understand the Data Vault concept, to be able to derive a good Data Vault model from source systems and to generate the appropriate ETL scripts.

Soon we'll launch version 1.0 of Quipu. With this product a source system oriented data warehouse can be constructed. Via a friendly user interface the user of the system is guided through consecutive logical steps in the process of designing, creating and maintaining of a technical and architectural robust data warehouse.

A source system model can be reverse engineered and will be stored in the repository of Quipu. Next an internal algorithm will apply derivation rules resulting in the proposal of a Data Vault storage model. Manually changes can be applied to this model after which this model also will be stored in the Quipu repository. Next standard mappings between source and target can be generated. These also will be stored in the repository. Also these mappings can be modified by the architect at will. Once ready Quipu deploys the Data Vault data warehouse model and accompanying load scripts.

A roadmap has been defined describing future releases of Quipu. These will include data mart generation (and ETL mappings). There are also ideas how to deal with the big 'T' or major transformation effort required between what we call the source Data Vault layer and the business Data Vault layer (the one containing more complex business logic). We will not develop an ETL tool. We will store the metadata of the major ETL vendors in the Quipu repository to support lineage and to allow companies to report every step in the data warehouse process flow. This is required because of increasing demands in the area of auditability and traceability.

We are convinced that this first release will help many companies in deploying an integration platform faster and cheaper. The first implementations have already proven themselves.

#### About Data Vault

The Data Vault Model is a flexible, scalable, and consistent "common foundational warehouse model architecture", invented by Daniel Linstedt. The modeling technique is covered by an implementation methodology. The Data Vault methodology is a set of data warehousing implementation components, consisting of project plans, design constructs, ETL loading templates, risk analysis, and repeatable, consistent build out strategies. The project / implementation methodology has its basis in SEI/CMMI Software engineering Level 5 components. More info can be found on [www.danlinstedt.com](http://www.danlinstedt.com) and [www.datavaultinstitute.com](http://www.datavaultinstitute.com). Interested in getting trained in the Data Vault methodology and modelling technique? [www.datavaultacademy.com](http://www.datavaultacademy.com)

**QOSQO**  
datawarehouse services

Karel Doormanlaan 1b, 5688 BP Oirschot (NL)  
Telefoon: +31 (0)499 577 562, Fax: +31 (0)499 577 059  
E-mail: [info@qosqo.nl](mailto:info@qosqo.nl), <http://www.qosqo.nl>  
Bank: ABN-AMRO 42.16.43.617  
K.v.K. Midden Brabant: 17226645, BTW 8194.65.847.B01