



Hello Readers,

Well our readership is expanding as more and more clients see that our newsletter is geared toward ERwin Data Modellers around the world.

Many thanks for the many data Modellers that came to our Toronto and Edmonton events in September.

As many folks know, CA World has just wrapped up. We know that a lot of ERwin users around the world could not attend for various reasons. However there were over 15 ERwin sessions that provided a lot of practical information for ERwin Modellers

We wanted to dedicate this issue to the highlights of CA World. There were a lot of very good ERwin sessions (Including 3 ERwin sessions from Sandhill), and we tried to cover them all.

We would like to give you our opinion on what we learned from this event. We will also provide a link for each presentation so that you could download it for review at a later time.

The next CA World will be in Los Angeles on April 11-15, 2010

As of this date (Dec 2008), the latest ERwin service patch V7.2.9 (Build 2458) is available for download. However, V7.3 is being released in the early part of the new year. We will let you know when this is available.

Robert Lutton,  
Editor, ERwin Times  
Sandhill Consultants



### ERwin – State of the Union

<http://caworld.com/Presentations/BIM06sna.pdf>  
<http://caworld.com/Presentations/BIM06sna.pdf>

Split across two sessions Sean Graham and Neil Buchwalter covered the current importance of data modelling, ERwin releases over the past year and plans up to the release of V8 of ERwin products, scheduled for 2010.

Sean spoke on the current importance of data modelling citing reports from Foresters, which identify Data Architecture as the hottest skill in IT. Neil spoke about what has happened in the ERwin suite over the last year and what is in the pipeline.

Looking forward we have two major events, the release of 7.3 which looks like it could be GA at the end of January 2009 and then Version 8 in 2010. There is likely to be a version 7.4 between these two.

Hot topics for 7.3 are comprehensive support for Teradata, a new integrated development environment ("IDE") for editing forward engineering templates, The new TLX macro language, new ODBC interface allowing SQL queries to be run directly against the ERwin model and a SQL Query tool.

A new member of the ERwin suite was also announced, ERwin Data Profiler from Exeros which is covered below.

Looking forward to Version 8 in 2010 we can expect major changes to Model Manager which will be supported on 5 platforms, comprehensive support for the DB2 DBMS family of products, exposure of the new Tom Sawyer graphics engine, property sheet and bulk update grids.

### CA ERwin – Integrates with Exeros X-Profiler for Data Profiling

#### CA World Announcement :

<http://www.ca.com/us/press/release.aspx?cid=194178>

If you have been reading our newsletters, you will notice that we have been starting to talk about Master Data Management. The reason is that CA has announced an integration with Exeros, as Exeros is a leader in the data relationship discovery area.

This integration provides another first for ERwin, the ability to offer comprehensive support for cross-system data profiling and analysis integrated with ERwin. Using both tools, data modellers who are concerned with data quality can now analyze the data in legacy systems, generate column level statistics and populate CA's ERwin Data Modeller with the data structure information for future modeling.

This powerful combination can also be used to perform cross-system analysis, comparing the data in any data source with other data sources, or with data in an existing data warehouse or master data management hub. The cross system analysis simplifies the integration of new data sources into a data warehouse, enables the consolidation of data marts into a single data warehouse, and provides the analysis required for creating and modeling a MDM hub.

Sandhill as a CA Partner are enabled to provide further information / demonstrations or pricing on CA's ERwin Data Profiler. We will also have a series of articles written for our Newsletter around ERwin and Data profiling in the coming quarters. In addition look for an upcoming webcast on this topic early in the new year.

For more information, please contact: [Robert.Lutton@sandhillconsultants.com](mailto:Robert.Lutton@sandhillconsultants.com)

## Sandhill – Complete Compare – Hands on Lab

<http://caworld.com/Presentations/BIM02LNA.pdf>

Two of Sandhill's Senior Consultants, Terry from the UK and Ed from New York, delivered two one hour lab sessions on complete compare.

This function remains one of the most powerful in ERwin and still holds a lot of interest for users.

The pair used a number of examples based on Sandhill's one day Complete Compare course and worked their way through a number of issues which users are presented with when using complete compare. The environment used for the workshops included different versions of the same model and a SQL Server 2005 database.

The sessions set up a number of "differences" between model and database and then examined these not just from the aspect of resolving them but to also examine how and why they arose in the first place.

Terry and Ed identified a number of key criteria which the user needs to be comfortable with to be comfortable and effective with this function. These included:

- Knowing how ERwin handles specific areas of data/functionality including
  - Attribute order
  - Column order
  - Database order
  - Constraints (Check, Default, PK, FK)
  - Unification
  - Reverse Engineering of objects into a model for comparison
  - Domains
- Managing comparison scope
  - Object types and properties
  - Object Instances
- Knowing how ERwin updates models

In working through their examples Ed and Terry took time to isolate each difference and cover all the issues related to it, from the change in the model or the particulars of a specific function which gave rise to it, through to either resolving it or changing the CC option set to ensure that the difference was not presented in future.

For more information on Sandhill Consultants ERwin Academy training program, please visit our website at [www.sandhillconsultants.com](http://www.sandhillconsultants.com) or e-mail [Robert.Lutton@sandhillconsultants.com](mailto:Robert.Lutton@sandhillconsultants.com) for more information.

## ERwin V7.3 – New IDE

<http://caworld.com/Presentations/BIM11SN.pdf>

Sandhill presented an introductory view of ERwin's new TLX Macro Language and the IDE where it can be employed to customise ERwin forward engineering templates. Both the language and the IDE will be released with ERwin V7.3.

Terry gave an overview of the syntax of the macro language and how it can be used to generate customised DDL. As an example he updated the FE template to selectively include a column in the DDL, the selection criteria reflecting whether the project adding the column had yet gone live. This example allows updates for multiple projects to be added to a model but only to be included in the DDL when the project itself goes live.

## ERwin and Microsoft's LINQ – Testing Triggers

<http://caworld.com/Presentations/BIM01SN.pdf>

Dr. James McCaffrey of Microsoft gave a presentation on the use of LINQ (Language Integrated Query) in the testing of triggers and the design of databases in ERwin which are LINQ compatible.

He showed how a test harness utilising LINQ technology could run a range of test cases against a database designed in ERwin and validate the results.

This approach can be used to make the testing of code objects such as triggers and stored procedures more systematic and thorough.

**SANDHILL**  
Your ERwin Partner of Choice

If you have a moment, check out our new website. We have updated our North American website to be dedicated entirely to CA's ERwin Data Modeling solutions. Sandhill's redesigned website focuses on the ERwin market and makes a great resource for the ERwin community.  
Feel free to visit our new N.A. website: [www.sandhillconsultants.com](http://www.sandhillconsultants.com)

### Metadata Integration

<http://caworld.com/Presentations/BIM07D N.pdf>

This presentation was delivered by Christian Bremau of Meta Integration Technology, Inc. The metaintegration bridges have been included in ERwin since release 7.1 and are now widely used by many companies.

One point of particular interest was the use of the bridges to assist the documentation of metadata from BI solutions, such as Cognos or Business Objects in relation to an Erwin model.

Christian demonstrated a number of scenarios where the use of the Metaintegration bridges to import and/or export meta can be used to great effect in managing BI solution meta data.

Particular areas he demonstrated included PK/FK Key analysis, Good snowflake design, Consistent logical naming Dimensional information

Christian was able to show how the use of the bridges can enhance the role of ERwin allowing a single model to become the repository for BI metadata rather than spreading it across many models and/or other tools.

Sandhill have found the Metaintegration bridges to useful in a number of scenarios and have experience of them predating their inclusion in ERwin Data Modeler. The extensive range of bridges included is impressive and is our first stop when dealing with data transfer between ERwin and other formats.

### Data Modeling for Governance

<http://caworld.com/Presentations/BIM09S N.pdf>

Unfortunately Sandhill did not have a chance to attend this session. However we have enclosed Marcie Barkin Goodwin's abstract on her presentation.

If you are engaged in data modeling efforts, whether large or small, you have a significant opportunity to produce models which are an integral part of a strategic and structured data governance effort. In a multi-user, multi-project environment, your modeling strategy should include the goals of data governance as well as 'best practices' modeling techniques for your tools of choice.

The strategy should drive the development of a structured approach, implemented by a supporting infrastructure. A repeatable, efficient modeling life cycle is critical to effectively managing the modeling environment, as well as implementing data governance requirements. Most organizations have a baseline set of standards for building logical and/or physical models, but lack critical 'integrity, reusability and redundancy' analysis, as well as a host of other necessary procedures to implement a strategic approach to modeling and governance.

A formal infrastructure is the only means by which modeling projects can consistently produce model structures with 'governance' integrity. This session discusses the importance of establishing a model management strategy which is designed and developed in conjunction with key data governance factors.

### Automating Changes using ERwin's API

<http://caworld.com/Presentations/BIM14S N.pdf>

This presentation by Kevin Arnold of EDS introduced the audience to ERwin's API and the flexibility it gives to updating and using ERwin data outside the tool.

Sandhill, who have long term experience of utilizing the ERwin API certainly support Kevin's message on the benefits of using the API.

Kevin spent some time on the structure of data in an ERwin model and how to identify the classes and properties of interest. ERwin V7.3 will ship with comprehensive documentation on the ERwin meta model adding another source of information.

He then covered the main elements of an API application with an example using Visual Basic for Applications ("VBA") within an Excel spreadsheet. The application he outlined would both extract data from an ERwin model into a spreadsheet, and use spreadsheet data to update an ERwin model.

Kevin introduced a nice twist with the potential for a library approach to increase flexibility. This approach may involve additional coding but its' offer of reuse should realise benefits downstream.

The library approach he advocated would typically be developed in a .Net language rather than VBA and he included some sample code using C#

### Data Management with ERwin

<http://caworld.com/Presentations/BIM16SN.pdf>

Antonio Amarin gave an interesting presentation which focused on the leverage of ERwin meta data in the wider task of data management.

He gave an outline of meta data within an organisation and ERwin's position as the host or logical, physical and RDBMS metadata within that space.

He also examined the different roles played by a number of tools within the ERwin suite, in relation to the management of metadata.

The ERwin suite members he identified specifically were ERwin Data modeller, ERwin Model Manager, The ERwin Saphir option and the new tool, ERwin Data Profiler.

As an example he spoke of a number of companies which had built a meta data process utilizing members of the ERwin suite and meta data repositories, some of which had been purchased and others which had been built in house.

### A Guide to implementing Industry Models

<http://caworld.com/Presentations/BIM17SN.pdf>

Tom Bilcze of the Westfield group gave a personal account of implementing an Industry model, in this case IBM's insurance industry models. He spoke of the need to be clear about what an industry model provides and of equal importance, what it does not! He cited the absence of a database design as an example of a "missing" component which had surprised some colleagues.

As the models are based on industry best practice they can provide real benefits and assist standardisation but terms may be too generic and will require customisation to reflect company specific data requirements.

These models also tend to be very large and working with them may need different approaches to those used with smaller models.

Sandhill's experience of working with industry models was similar to Tom's. They offer benefits, but to get the most out of them requires an informed and considered approach.

## Automatically Analyze Your Legacy Data to Populate CA ERwin Data Modeler Data Models!

<http://caworld.com/Presentations/BIM12SN.pdf>

This session was entertaining as well as informative. If you ever have the chance to see Todd Goldman present you will find his presentations very enjoyable.

As you know, CA has inked a deal with Exeros to integrate with ERwin Data Modeler. The issues that many data modellers face are that some don't know what is in their data sources, or perhaps some ERwin data models are not as up to date as they should be for all data sources.

CA's ERwin Data Profiler will allow you to understand your current state. It can drive legacy models, determine the quality of the model and match the consistency of different models. On top of all that, ERwin Data Profiler can validate that the data in your data sources matches the meta data of your ERwin models.

Todd demonstrated some of the standard points within a data profiling solution. Such as, generating a bunch of stats on the data for cleansing, understanding patterns within the data and a drill down to review the details of the data. However, there are a couple of items that make CA's ERwin Data Profiler unique. It's focus on the identification of the primary foreign key analysis; all other tools leave you to figure this out. In addition it can cross profile up to 20 different data sources, and then allows you to drill down and identify the duplications and show any orphan rows (or Dirty Data, as it is referred to).

When you add the fact that it does all of the above and provides an overlay analysis of all the data sources in one spot, you can see that it won't take long before CA ERwin data modellers are asking to have this as part of their tool kit to design and implement great data bases, for either a Master Data Management or Data Warehouse projects.

When asked for a couple of examples, the following use cases were giving.

#1. CA's ERwin Data Profiler could be used for data profiling, specifically to compare the metadata of an ERwin data model vs the instance data in the database are the same. Apparently in some databases, Modellers would sometimes use an available column for a purpose other than what it was intended for (Go figure!!).

#2. The second use would be the ability to import from a flat file, IDMS, Copy Book or Excel into ERwin to give you a first cut of the Data model.

#3. Some data modellers talked about using CA ERwin Data Profiler to R.E. a DBMS import into ERwin and then do a complete compare with the ERwin Data Model. The difference between this approach and just R.E. into ERwin directly is that ERwin Data Profiler will do a great job analysing the primary foreign keys.

There were several other options that were covered and I would recommend anyone who is interested in Data Quality to take a look at the PPT and if you are interest in more information, please contact [Robert.Lutton@sandhillconsultants.com](mailto:Robert.Lutton@sandhillconsultants.com)

## Teradata – Managing Extremely Large Models using ERwin.

<http://caworld.com/Presentations/BIM08SN.pdf>

Teradata ERwin models are large. Most models have up to 2,100 entities with 8,800 attributes and are in the range of 30MB in size. What I found interesting is that they are currently at v4.x and are waiting until V7.3 (which has full Teradata support) is released before they consider upgrading. When e-mailing files they split the files and then use Winzip to send them.

Teradata have 8 industry type models, (Financial Services, Health Care, Manufacturing, Communications, Media and Entertainment, Travel, Retail and Transportation). They have one Major Data Architect per industry model and they save their models every 30 minutes. Their models all have a conceptual and logical model. All models are normalized to 3NF.

Teradata use API scripts for many different things including managing Teradata files and models. They also make use of UDP's for unique numbering of Teradata objects.

The way Teradata perform data modeling they do not need to use Model Manager. One thing they pointed out was that their text boxes, which are all over the model, cluttered up Model Manager. Some Teradata clients, however, do use Model Manager and they tend to have a master model and then create smaller subject area models from their master. (So for Teradata users who are using model manager watch out for text boxes).

Teradata enforce standards through their models and they also make extensive use of color for managing and moving attributes around (currently they have 83 different colors in use). They use a lot of the standard ERwin reporting solutions. What they tend to do re documentation is to generate via ERwin and then bring it into a Word document for additional enhancements.

Teradata are using ERwin to build other tools / solutions, one example being an Enterprise Data Warehouse roadmap for their solutions.



Sandhill's next event will be held at the Canadian War Museum in Ottawa, Canada on January 20<sup>th</sup>, 2009.

This major national museum of military history, covering Canada's military past from earliest times to present day, is the venue of choice for Sandhills latest ERwin Event.

For registration please e-mail [Robert.lutton@sandillconsultants.com](mailto:Robert.lutton@sandillconsultants.com)