Satellite Navigation for Meta data
(Discovering Information Structures)
Hannah Roberts
Session Code ET02
Hannah Roberts

— Data Analyst with over 30 years IT experience
— 9 of the last 10 as a Data Consultant for a large UK Financial Institution
— Lead a team which established an enterprise-wide approach to data architecture and data governance
Finding and Navigating through large databases

- Is a challenging task: complex, time consuming, costly, particularly where current documentation is not easily available
- Is a fundamental part of integration/ migration projects, but also warehouse builds, master data management projects, etc.
- with project and business risks if not correctly completed

This presentation introduces a structured approach which

- will be particularly helpful when exploring ERP systems
- gives simple systematic steps to discover the database structure
- provides a more robust foundation for complex data related projects
- speed up the data discovery process where some tools can be used
- suggest some potential tools to support the method
Agenda

- Context
- Outline the Discovery process
- Process walkthrough using Saphir 5.5 & Erwin 8
Integration / migration projects need to identify:

- What table information to move from one system to another
- Whether the same concept exists in both systems?
- What the database names are of the particular tables and fields

Common task also needed for:

- Building data warehouses
- Implementing master data management
- System Instance rationalisation
What information do we need for this task:

- Table names and definitions
- Column names and definitions
- Primary keys and Foreign keys

Where can we get this information:

- Database catalogue - reverse engineer with ERwin
- Model documentation - scanning through textual information
- Data Dictionary particularly for ERPs – some supplier tools or Saphir
Large databases

- Contain thousands of tables, millions of fields
- Names not necessarily meaningful in the database catalogue
- Documentation may not be easily accessible or reliable

This is then becomes a more challenging task?
Where is this a particular challenge?

- Enterprise systems, e.g. SAP, PeopleSoft, Oracle EBS, Siebel,
- Complex internal databases

Who needs to understand the challenge?:

- Business Analysts, data architects, information specialists
- Data Managers, Business Strategists, Process Architects
The Discovery Process

Some basic database constructs

- Table - holds records of information
- Relationship - links a record in one table to a record(s) in another table

Allows functions to navigate through the database structure

- Many functions using a table - accessed via different relationships
The Discovery Process
Determine the start point
Identify the related tables

Determine the start point
Identify the related tables
Select manageable sets to analyse

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Omit tables not in scope

Omit tables not in scope
Validate the emerging structure

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Omit tables not in scope
Validate the emerging structure
Explore the sub levels of information

- Determine the start point
- Identify the related tables
- Select manageable sets to analyse
- Check any textual descriptions
- Omit tables not in scope
- Validate the emerging structure
- Explore the sub levels of information
Repeat the exploration for other start points

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Omit tables not in scope
Validate the emerging structure
Explore the sub levels of information
Repeat the exploration for other start points
Release the model for publication

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Omit tables not in scope
Validate the emerging structure
Explore the sub levels of information
Release the model for publication
Discovery Demonstration

Tools

- Saphir 5.5
- ERwin 8

Scenario

- A company wants to outsource some of its order packaging.
- An export is needed from SAP of the information about orders
- What SAP tables and fields should be used?
The DISCOVERY process - Summary

Determine the start point
Identify the related tables
Select manageable sets to analyse
Check any textual descriptions
Omit tables not in scope
Validate the emerging structure
Explore the sub levels of information
Release the model for publication
The Discovery Process - Questions

Determine the start point

Identify the related tables

Select manageable sets to analyse

Check any textual descriptions

Omit tables not in scope

Validate the emerging structure

Explore the sub levels of information

Release the model for publication
Thank You

Add contact information:

Name: Hannah Roberts  
Phone: +44 (0)1344 638102  
Email: hroberts@silwoodtechnology.com  
Website: www.silwoodtechnology.com
© Copyright Silwood Technology. All rights reserved. All trademarks, trade names, service marks and logos referenced herein belong to their respective companies. No unauthorized use, copying or distribution permitted.

THIS PRESENTATION IS FOR YOUR INFORMATIONAL PURPOSES ONLY. CA assumes no responsibility for the accuracy or completeness of the information. TO THE EXTENT PERMITTED BY APPLICABLE LAW, CA PROVIDES THIS DOCUMENT “AS IS” WITHOUT WARRANTY OF ANY KIND, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT. In no event will CA be liable for any loss or damage, direct or indirect, in connection with this presentation, including, without limitation, lost profits, lost investment, business interruption, goodwill, or lost data, even if CA is expressly advised of the possibility of such damages.

Certain information in this presentation may outline CA’s general product direction. This presentation shall not serve to (i) affect the rights and/or obligations of CA or its licensees under any existing or future written license agreement or services agreement relating to any CA software product; or (ii) amend any product documentation or specifications for any CA software product. The development, release and timing of any features or functionality described in this presentation remain at CA’s sole discretion.

Notwithstanding anything in this presentation to the contrary, upon the general availability of any future CA product release referenced in this presentation, CA may make such release available (i) for sale to new licensees of such product; and (ii) in the form of a regularly scheduled major product release. Such releases may be made available to current licensees of such product who are current subscribers to CA maintenance and support on a when and if-available basis.