Documenting ETL Rules using CA Erwin Data Modeler

By Sampath Kumar

Abstract

In any data warehouse development project some of the major challenges include

- Effective capture and maintenance of metadata information in data model such as data source ,transformation rules and data synchronization etc
- Effective communication of captured metadata information by data modeler to other teams such as ETL

This document covers features in CA Erwin Data Modeler which can be leveraged for capturing the metadata information such as Extract Transform Load (ETL) rules. This document explains step by step of how to capture the ETL information using Erwin and also covers the generation of reports with the captured information to communicate effectively to other teams.

Introduction

The data warehouse combines information from several **Online Transactional Processing** (OLTP) systems and archive data into a single decision support system. It can be either relational or non relational data source (both structured and unstructured data). In order to keep the data in synch with the operational system it's very essential to capture the data source for each column in the data warehouse and information of when and how the data is updated.

So in a nutshell the following information needs to be captured in any data warehouse environment

- Source of data
- Transformation rules-The method in which the data is getting extracted, transformed and loaded
- Frequency: The frequency and timing of data warehouse updates.

In many organizations it used to be a separate document apart from data model but it becomes very hard to maintain document and data model in synch.

Why it's important

Data modeling is the first step which converts the business rules into a data model and the data modeler is the one who understand the rules from business counterparts (both in a structured and unstructured way). As a result of this, the data modeler captures most of the business rules directly in the data model and some of them (such as data source, transformation rules and frequency above) needs to be passed on to other teams such as DBA and ETL .It's very essential to capture all data related business rules as a part of data modeling effort to avoid getting lost. CA Erwin Data Modeler provides effective way to capture this information and as a data modeler it should be captured as part of modeling efforts.

Approach

This document covers how the above challenges can be addressed using Data Transformation and Data Movement features available in CA Erwin Data Modeler. To explain better there will be a simple running example throughout this document which will navigate step by step.

Overview

The CA Erwin Data Modeler has come up with the following salient features to capture the metadata information effectively.

- Data warehouse Sources Dialog: to define sources of data for your data warehouse
- Columns Editor: to document the data warehouse source assignments and transform the information for each column in the dimensional model in the data source tab.
- Data Movement Rules Editor: to document the data warehouse maintenance processes required to regularly update each table in your dimensional model.

Let's explore these features in detail in the rest of the document using simple example of Customer_Dim.

Customer_Dim

Let's take a fictitious example of an entity Customer Dimension to explain the above features. Let's assume that it's sourced from multiple operational systems (relational DB), attributes having different transformation comments and the frequency of customer information getting updated is daily.

The following are the key attributes in the Customer dimension

- Snapshot
 - customer_SKID
 - snapshot_Begin_Date
 - snapshot_End_Date
 - current_ind
- Basic Information
 - o Customer name
 - Customer Date of Birth
 - o Driving License
- Address
 - Mailing Address
 - Physical Address
- Communication
 - o Email Address
 - o Phone
 - ∘ Fax
 - Segmentation
 - Shopping
 - o Behavior

Capturing Data Source

In order to proceed further lets create the empty data model using "Create Model" (File ->New) of the model type Logical/Physical and target database as Teradata

Create Model - Sele	ect Template		×
New Model Type	O Physical	• Logi <u>c</u> al/Physical	OK
Create Using Templa	te: ical Model		
Remove	Browse <u>F</u> ile System	Browse AllFusion MM	
Creates a new mode classic) and default (l with both logical and physic settings.	al levels (AllFusion ERwin DM	
Target Database			
<u>D</u> atabase: Te	radata 🔽 🔽 Vers	sion: 2x 💌	

Let's create the Customer_Dim table and add the attributes

🗣 AllFusion ERwin Data Modeler by CA - [Customer_Dim.erwin : <main area="" subject="">]</main>	- 2 🗷
님 Elle Edit View Format Model Database Tools Services Window Help	_ 8 ×
🛅 🙆 🖺 👌 攝 驫 🚯 🏦 🝄 屆 🗟 🌑 🔍 🍳 🔍 🔍 🎕 🔯 🕞 ± Physical 🕑	
[🔍 🗃 🕞 노 노 🔤 Microsoft Sans Serif 🔽 10 💌 🖪 🖉 ± 조 ± 소 💷 超 밝 湘) 光 옷 🗮	i 🗮 📴 🗗 🔺 🔍 🖼 🕼 🗢 🛝 😢 💐 🕸
<u>■□□○/\$₹∎</u> [[葉古☆☆□日☆☆★ ● ●] ♥	
Eustomer_Dim customer_skid snapshot_edi_date snapshot_end_date current_ndr customer_nbr first_name last_name last_name mailing_address_line1 mailing_county mailing_county mailing_state physical_address_licat_nont physical_county physical_county physical_county physical_county physical_county physical_county physical_county physical_county physical_state physical_state physical_state physical_segment_nbr behavioral_segment_nbr behavioral_segment_nbr shopping_segment_br shopping_segment_desc	
Display1	
Davids	LIGHEOFFICE June 20

In order to make this model as dimensional model and to capture the data movement rules **goto Model->Model Properties** and select the check box for Dimensional and Data Movement



Model Properties
General Definition Notation Defaults RI Defaults UDP History Options History Model Info Name: Customer_Demographics Author: Type: Logical/Physical Database: Teradata Enable Modeling Features Optional Optional Optional Show source objects in logical, target objects in physical Auto apply Many-to-Many transform Auto apply Supertype-Subtype Identity transform
OK Cancel

You can observe that in the columns wizard "Data Source" tab will be enabled since the "Data Movement" was selected.

Columns	× Normal State Sta
Table: Customer_Dim	💌 🔜
Column Column Column Customer_skid snapshot_begin_date snapshot_end_date current_ind customer_nbr first_name: String/VARCHAR(20) last_name mailing_address_line1 mailing_address_line2 mailing_county mailing_city mailing_city mailing_state New Rename Delete	Constraint Comment UDF Data Sources:* Data Sources:* Customer_Operational_DB.Customer.customer_first Transform Comment:* Description (Description) Type 2
	OK Cancel

Click on the button (with 3 dots) to open the "Data sources" wizard to create the data source which can be used for mapping the columns

Data Warehouse Source Selector	\mathbf{X}
Available Sources:	
	Select
	Edit
Selected Data Sources:	
	Deselect
	ОК
	Cancel

Click the edit button of "Data Warehouse Source Selector", to open the below screen. In the source information, provide the Operational Source System name, operational database host name, operational DBMS type. Select "Type" as Relational if the operational system is relational or flat file if it's from the file feed.

Data Warehouse S	iources 🔀
Source Name	
Customer_Operation	nal_DB
General Detail	efinition
-Source Information	ion
System Name:	Customer_Ops
Host Name:	Customer_Ops_Hostname
DBMS Type:	DB2 UDB Version: DB2/UDB 8.1
	Type Relational O Flat File
Imported From:	
	OK Cancel

Open the tab "Detail" which will provide the options of "Import from CSV" and "Import other". Click the "Import Other"

Data Warehouse Sources			×
Source Name Customer_Operational_DB			
General Detail Definition	Columns:		
	Name	Datatype	Comment
New Rename Delete		Import fro	om .CSV File
			OK Cancel

The "Import other" provides three options to import the table structure

- Flat File
- Database/Script
- Model Manager

Data Source Import			X
The Right Model Load or select a model from whic	h Data Source Objects ar	Compare Leve re to imported.	el: Physical
	-Load From		
	◯ File	 Database / Script 	
<u>Overview</u>		Allow Demand Loading	Load
Current Model	Open Models in Memor	ry:	
	Open Models	Location	
The Right Model	Custom Report S	tr ModelMart://modelmart/Z - HC	LD/customer opti
Type Selection	Customer_Demo	gr <unknown></unknown>	
<u>Type selection</u>	Model_1(1)	ModelMart://modelmart/DATA	WAREHOUSE/Cor
Left Object Selection	Model_1(2)	<unknown></unknown>	
	Model_2	<unknown></unknown>	
Advanced Options	Model_5	ModelMart://modelmart/Z - HC	DLD/Customer
	< <u>B</u> ack <u>N</u> ext	t > Import Clos	e Help

In our example let's assume that we are importing the table structure from the "Model Manager" .Select "**Model Manager**" and click on "**Load**" button open the Model Mart to load the model from the operation system. Let's assume the following is the high level model of Customer data in the operational system and we want to map these tables to dimensions in data warehouse.



Click on "**Import**" to import the data model from the model mart which will display all tables in that data model .Select the Customer, Customer_Address, Customer_Segmentation, Behavioral_Segment and Shopping_Segment



Click on the small arrow (highlighted in red circle) in the above diagram to select the required tables from the model mart. Once the required operational system tables are selected; it will be available for mapping as below.

Source Name Customer_Operational_DB General Detail Definition Tables: Behavioral_Segment Customer_Address Customer_Address Customer_Segmentation Shopping_Segment Outromer_Segmentation Shopping_Segment VarcHAR(Outromer_Segmentation Shopping_Segment Outromer_Segmentation Shopping_Segment Delete Import from .CSV File	Data Warehouse Sources				X
General Detail Definition Tables: Behavioral_Segment Customer_Customer_Address Customer_Segmentation Shopping_Segment Shopping_Segment VarcHAR(Outmos: Customer_last_name VARCHAR(Outmos: Customer_last_name VARCHAR() Outmos:	Source Name Customer_Operational_DB				
Derivitival_segment Name Datatype Commel Customer Customer_Address INTEGER INTEGER Customer_Segmentation Shopping_Segment VARCHAR(Image: Commel and the segment	General Detail Definition	Columns:			
Inviring_incense_ind Inviring_incense_ind driving_license_state CHAR(2) driving_license_nbr INTEGER driving license state CHAR(2) Viring_license state CHAR(2) Import from .CSV File Import Other	Customer Customer_Address Customer_Segmentation Shopping_Segment	Name [customer_id customer_first_name customer_last_name dob diving_license_phr	Datatype INTEGER VARCHAR(VARCHAR(date INTEGER	Comme	
New Rename Delete Import from .CSV File Import Other		driving_license_ribi driving_license_state driving_license_nbr driving_license_state	CHAR(2) INTEGER CHAR(2)	>	
	New Rename Delete	Import from	.CSV File Imp	ort Other	

Click "Ok" to see the source tables available for selection.



Columns	$\overline{\mathbf{X}}$
Table: Customer_Dim	~
Column Colored State Sta	Constraint Comment UDP Data Sou Data Sources:* Customer_Operational_DB.Customer.customer_first Customer_Operational_DB.Customer.customer_first Transform Comment:* Type 2 OK Cancel

Once all the "Data Source" and "Transformation Comments" are entered, ETL Spreadsheet can be generated .Please follow the below steps to generate the ETL Spreadsheet.

- 1. Highlight the tables for which you want generate the ETL spreadsheet
- 2. Click on the "Data Browser" icon which will open the Data browser window for all possible reports which can be generated.

Create new template "ETL Spreadsheet.erp" report using "Data Browser".



Open the "Reports" in the menu and select "New Reports File" .Create the report named "ETL Spreadsheet.erp"



Save As						? 🔀
Save <u>i</u> n:	C Reports		~	G 🦻	بي 🥙	
My Recent Documents Desktop My Documents	HTML Table					
My Computer						
	File <u>n</u> ame:	ETL Spreadsheet			~	<u>S</u> ave
My Network	Save as type:	Report Files (*.erp)			~	Cancel

Right click the newly generated report click Edit ERFwin Report 'ETL Spreadsheet'

🔕 Data Browser					_ 2 🔀
File Edit Search View Reports Help					
🖀 🧉 🗟 💋 Table Reports : ETL Spreadsheet 🛛 💌 🕨 📲	16 8 8 12 ← →				
All reports	Result Set				
E-G General					
E- Diagram Benorts					
Model Validation Reports					
🖻 😋 Other Reports					
Diagram Contents by Class					
B7 Non loadable versions					
E G Action Summary Reports					
Session Summary Report					
Session Summary - Detail Report					
Compare original version of wode_5 with current version of wode_5					
Physical Object Calculations					
- Zatabase Object Calculations					
AllEusion EBwin DM Benorts IC:\Program Eiles\CA\AllEusion EBwin Data Mr.					
E 🔄 Table Reports					
ETL Spread Edit ERwin Report ' ETL Spreadsheet'					
Execute report 'ETL Spreadsheet'					
Rename 'ETL Spreadsheet'					
💠 4m fan 🎾 🗙	<	III			>
🛃 start 📄 😸 🗿 🖾 🐻 💯 🕼 🧎 😭 C:\Samp 🔛 D	ocumen	CA Clarit Remedy	🗿 Inbox 🛛 🖂 RE: DA R	app (Rea	🗿 😒 🕙 💟 🕋 📳 5:57 AM

Make sure only the following options are selected which is relevant for ETL.

Reports	
Reports Name: ETL Spreadsheet Cat Options Definition Note Options: Image: Column Image: Column Image: Column Image: Column Image: Column <t< td=""><td>○ Logical ○ Physical egory: Table ▼ Options ○ Edit Show ⊙ Selected Only</td></t<>	○ Logical ○ Physical egory: Table ▼ Options ○ Edit Show ⊙ Selected Only
	Show Selected Collapse All Clear All OK Cancel

Save the report to make sure selected columns are stored in the report template.





Expand "CA Erwin DM Reports" ->Expand "Table Reports" ->Expand "Physical only Table/Columns" ->Double click on the "ETL Spreadsheet" which will generate new report

🔕 Data Browser							_ 2 🛛						
File Edit Search View Reports Help													
🖀 🗃 🛕 🚳 Marchard Columns : ETL Spread	dshei 📉 🕨	• * * * * •	⇔ =	0 E 🗉	60°								
All reports	Physical Only T	able/Columns : ETL Spread	sheet	(C:\Sampath\Offi	cial\CA Erwin\Customer_[Dim.erwin, 10:20:09 AM, 23 rows)							
🖽 💼 General	🧷 Table Name	🧷 Column	Key	🧷 Datatype	🧷 Transform Comment	Source Name	Source Datatype						
😑 😁 AllFusion ERwin DM Reports [C:\Program Files\CA\AllFu	Customer_Dim	customer_skid	PK	INTEGER	Surrogate		1						
Attribute Reports		snapshot_begin_date		INTEGER	Snapshot								
🕀 🧰 Column Reports		snapshot_end_date		DATE	Snapshot								
Entity Reports		current_ind	411	UHAH(1)	Current	Culture Republical DB Culture submer id	INTEGER						
E- Comain Benorts		first name	AKT	MADCUAD(20)	Tupe 2	Customer_Operational_DB.Customer.customer_id	VAPCUAP(20)						
E-B Table Benots		uis(_ridine		VALICI MI (20)	Write your additional	customer_operational_ob.customer.customer_msr_name	VALICIAL(20)						
Table/Owner					transformation								
Table/Comment					comments)								
Tabler Commerk		last_name		VARCHAR(20)	Type 2	Customer_Operational_DB.Customer.customer_last_name	VARCHAR(20)						
					Write your additional								
E-12 Envisical Univ Fability Selumns					transformation								
Ender ETL Spreadsheet		mailing address line1		VARCHAR(20)	Commentsj Tune 2	Customer Operational DR Customer Address mailing address line1	VARCHAR(20)						
B ETL Spreadwheet (C:\Sampath\Official\		mainig_address_inter		White Milling	Write your additional	customer_operational_brb.customer_wadress.mailing_address_inter	VALICI MI (20)						
ETL Spreadsheet (C:\Sampath\Official\					transformation		=						
ETL Spreadsheet (C:\Sampath\Official\					comments)								
ETL Spreadsheet (C:\Sampath\Official\		mailing_address_line2		VARCHAR(20)	Type 2	Customer_Operational_DB.Customer_Address.mailing_address_line2	VARCHAR(20)						
ETL Spreadsheet (C:\Sampath\Official\					(Write your additional								
B ETL Spreadsheet (C\Sampath\Official\)					transformation								
		water and here	VARCHAR(20)	Commentsj	Customer Reportional DR Customer Address mailing sounds	MARCHAR(20)							
E DPA Papart		maining_county		VANCHAR(20)	Write your additional	customer_operationar_bb.customer_Address.maiing_county	VANCHAN(20)						
E Critical Area Denaits								transformation					
Em Subject Area heports													
H Model Validation Reports		mailing_city		VARCHAR(20)	Type 2	Customer_Operational_DB.Customer_Address.mailing_city	VARCHAR(20)						
End Stored Procedure Reports					(Write your additional								
🗄 🗀 Relationship Reports				transformation									
E- Carlo View Reports				CHAD(2)	comments)	Customer Reportional DB Customer Address maline state	CUAD(2)						
🗄 🗀 Action Summary Reports		maiing_state		UHAH(2)	1 ype 2 04/rite upur additional	Customer_Operational_DB.Customer_Address.mailing_state	UHAH(2)						
🗄 🦕 Volume Reports					(write your additional								
H- AlEusion MM Benots (modelmait)					comments)								
- •		mailing country		CHAB(2)	Type 2	Customer Operational DB.Customer Address.mailing.country	CHAB(2)						
					Write your additional								
					transformation								
					comments)								
		physical_address_line1		VARCHAR(20)	Type 2	Customer_Operational_DB.Customer_Address.physical_address_line1	VARCHAR(20)						
					Write your additional								
					(ransroimation								
		physical address line?		VARCHAR(20)	Tune 2	Customer Operational DB Customer Address physical address line?	VABCHAB(20)						
		prijolog_dddroos_moz		(Farter Fartied)	Write your additional	carana_operatoria_op.caranat_radicat.prijstod_datros_inter	The fact the factory						
					transformation								
					comments)								
		physical_county		VARCHAR(20)	Type 2	Customer_Operational_DB.Customer_Address.physical_county	VARCHAR(20)						
<					(Write your additional								
					transformation		×						
💠 hao hao 🎾 🗙					comments								
	1		_				1						
🥑 start 🚽 🤅 🗿 🗷 🗟 📨 🧭 🕴 🔘 in	🖂 d	🌒 R 👔 I		t 🖥 /	l 🙆 Ci 🦉] D 및 All 뭘 data 뭘 Sa 뭘 ET	🧿 🛃 🍳 😋 🗗 🕅 10:36 AM						

Right click again and select "Export result set ETL Spreadsheet" which will open the "Export from Data Browser" window .Select "CSV" in the Export

Export from Data Browser									
ETL Spreadsheet (C:\Sampath\Official\CA Erwin\Customer_Dim.erwin, 10:20:09 AM, 23 rows)									
Export CSV Convert the result set into a CS values) file. The output can be i spreadsheet applications, such	Export CSV Comma separated values) file. The output can be imported by most spreadsheet applications, such as Excel.								
Presentation Tabular Tabular with duplicates Master-detail Indented Item per line	Export to File Clipboard Export headings								



The final ETL Spreadsheet will look like the following which will be used as deliverable to the ETL team. You can also highlight the important details after generating the report from the tool like the following .Save it as .**xis** type so that all your custom made changes will be retained when you open again.

× 1	Nicrosoft Excel -	ETL_Spreadsheet											_ 7 🔀
: 🖻	<u>F</u> ile <u>E</u> dit <u>V</u> iew	Insert Format Tools D	ata Te	<u>a</u> m <u>W</u> indow <u>H</u> e	elp						Type a ques	tion for help	8×
	📂 🛃 🖪 🔒	🗃 💁 🖤 🛍 👗 🖻	🛍 - 🥥	1 10 - 12 - 1	🧟 Σ - Ζ↓ Ζ↓ 🛍 🦏	0	🙄 Arial	- 10 - B <i>I</i>	Ū ≣ ≣	≣ • a • \$	% , .0 .00	🗐 🖽 🗸 🗳	» - <u>A</u>
		M 17 M 18 10	₩ Reply	with Changes E	nd Review								
: ***	New List Get \	Nork Items 💼 Publish 🔳 Refi	resh E	Configure List	Choose Columns Links a	and Attac	chments						
	G24 -	★ VARCHAR(20)					F						
	A	В	С	D	E			F			G	Н	I 🔨
1	Table Name	Column	Key	Datatype	Transform Comm	ent	Source Name				Source Datatype		
2	Customer_Dim	customer_skid	PK	INTEGER	Surrogate			NA					
3		snapshot_begin_date	AK1.2	INTEGER	Snapshot			NA					
4		snapshot_end_date		DATE	Snapshot			NA					
5		current_ind		CHAR(1)	Current			NA					
6		customer_nbr	AK1.1	INTEGER	Fixed		Customer_Operational_DE	3.Customer.custom	ner_id		INTEGER		≡
					Type 2								
_		_			(Write your additional				_				
7		first_name		VARCHAR(20)	transformation commen	ts)	Customer_Operational_DE	3.Customer.custom	ner_first_name		VARCHAR(20)		
					Type 2								
8		last name			transformation commen	te)	Customer Operational DF	3 Customer custom	ner last name				
- U		last_hame			Type 2	1.5)	Customer_operational_bt	5.Gustomer.custon	ner_last_hame				
					(Write your additional								
9		mailing_address_line1		VARCHAR(20)	transformation commen	ts)	Customer_Operational_DE	3.Customer_Addres	ss.mailing_add	ress_line1	VARCHAR(20)		
					Type 2								
					(Write your additional								
10		mailing_address_line2		VARCHAR(20)	transformation commen	ts)	Customer_Operational_DE	3.Customer_Addres	ss.mailing_add	ress_line2	VARCHAR(20)		
					Type 2								
11		mailing county			(VVrite your additional	to)	Customer Operational D	Customor Addrog	oo mailing	entre i			
		mailing_county		VARCHAR(20)	Type 2	15)	Customer_Operational_DE	S.Gustomer_Addres	ss.maiing_cou	inty	VARCHAR(20)		+
					Write your additional								
12		mailing city		VARCHAR(20)	transformation commen	ts)	Customer Operational DE	3.Customer Addres	ss.mailing city	,	VARCHAR(20)		
					Type 2								
					(Write your additional								
13		mailing_state		CHAR(2)	transformation commen	ts)	Customer_Operational_DE	3.Customer_Addres	ss.mailing_sta	te	CHAR(2)		
					Туре 2								
14		mailing country			(Write your additional	ta)	Customer Operational DE	Quatamar Addres	oo mailing .co:	entre (
14		maning_country		CHAR(2)	Type 2	is)	Customer_Operational_DE	5.Customer_Addres	ss.mailing_cou	nuy			~
H 4	(► ► ► ETL_Spr	eadsheet/						<		1111			>
Read	dy												
2	start	o 🛛 🖬 🗿 😿 🕑 👌	🕒 Inb	. 🖂 dat	🗿 Re 👔 IT	🔁 Th	n 🛛 🏠 C:\ 🖉 🕎 Do	. 🖫 All	월 data 🛛 🖉	Sa	l) et 🖂 🙆 🥩	® 🕒 🗗 🕅	11:02 AM

Data Movement Rules in Erwin

Data movement rules in the Erwin Data Modeler enable you to maintain processes required to regularly update all tables in the model. In our example it's used for keeping the Data warehouse and the operational source system in synch, the various management rules used to manage the information supported by Erwin are:

- Refresh: Replaces existing data.
- Append: Updates the existing information with changes and additional information.
- Backup: Creates a copy of the information to make it available for recovery.
- Recovery: Stores information from the backup information, the recovery process is required when the data is
 lost due to hardware or network failure.
- Archiving: Extracts information from tables based on criteria and saves the information in a file for future reference.
- Purge: Extracts information based on criteria but does not save the information.

Open "Model" in the main menu and select "Data Movement Rules"



In the "Data Movement Rules" window define the rule name as "Snapshot" and type as "Append" (Refresh, Append, Backup, Recovery, Archive and Purge) .In the definition tab explain the meaning of the rule and how it needs to be attached to tables in the Data warehouse.

ta Movement Rules		
	Delete	Type Filter: All
Rule Name		Туре
Snapshot		Append
Definition Attachment Definition:	ening in the source syste	m and it will be appended to
		×
		OK Cancel

The "Type" is a drop down and it will have following options.

Refresh
Append
Backup
Recovery
Archiving
Purge

Once data movement rules are defined, attach the rule to the table

Teradata Tab	es 🛛 🔀
Table: Custom	er_Dim
Name: Dustome	er_Dim Database:
Dimensional	Data Movement Physical Props Comment Volumetrics UDP Hi] < 🔊
Data Movem	ient Rules
<u>R</u> efresh:	
<u>A</u> ppend:	Snapshot 💌 📖
Backup:	
Reco <u>v</u> ery:	
Arc <u>h</u> iving:	
Purge:	
Physical Only	v ✓ <u>G</u> enerate DB Sync OK Cancel

In our Customer_Dim we want to capture the changes happening in the source system as Type 2 dimension so we have selected the "Append" rule named Snapshot. For Type 1 dimensions use "Refresh" rule.

To retrieve this data in the form of report goto "Data Browser" select "File-New Report" .It should be generated separately apart from the "ETL Spreadsheet" to capture the table level information.

Reports		$\mathbf{\overline{N}}$
Report <u>N</u> ame: Data Movement Rule	O Logical ⊙ Physical Category: Table	
Options Definition Note		
	 ○ Petions ○ Edit ○ Show ○ Selected ○ Only Show Selected ○ Collapse All ○ Clear All 	OK Cancel

Export the report in the form spreadsheet

×	Microsoft Excel	- Data_Moven	nent_Rules							- 7 🔀
: 3	Eile Edit Viev	w <u>I</u> nsert F <u>o</u> r	mat <u>T</u> ools	Data Team Windo	w <u>H</u> elp				Type a question for help	8×
10	i 💕 🖬 🖪 🔒) 🖪 🖪 🖤	' 🛍 i 🐰 🖻	🖻 🔁 = 🛷 🔄 = (🖻 🗸 🕵 🗴 🗕 🛓 🗸 🕹 🛍 🦓 🔞 📲 🔡 Ari	al	• 10 • B I U ≣ ≣ ≣	≣ 🔤 \$ % ,		3 - A -
18	1 fa fa 🖉 👒	100 B) 🌒 🖏 😥	Reply with Chang	es End Review					
1										
	F12 V	fy								
	A	B	С	D	E	F	G	Н	1	J
1	Table Name	Refresh Rule	Refresh D	Append Rule Nan	Append Rule Definition	Used by Table	Used by Table Comment	Backup Rule Nam	Recovery Rule Nam	Archive R
2	Customer Dim			Spanshot	This rule is used for capturing the changes happening in the source system and it will be capacided to the civiting data in the table.	Customer Dim	The Customer_Dim is a Type 2 dimension which will capture the changes happening in the source system and snaps the			
2	Customer_Dim			Snapshot	appended to the existing data in the table.	Customer_Dim	changes.			
4										
5										
6										
7										
8 Q										
10	-									
11										
12						1				=
13										
14										
15										
17										
18										
19										
20										
21										
22										
23										
25										
26										
27										
28										
29							1			×
H	♦ ► ► ► ► ► ► ► ► ► ► ► ► ► ► ► ► ► ► ►	lovement_Ru	iles/				<			>
Rea	dy									

Conclusion

The metadata information such as "Data Source", "Transformations rules" and "Data Movement rules" are very important for any Data warehousing efforts and it's very critical to capture the correct information. It will be the guideline for the ETL team to create mappings to source system and load as per the rules. So try to provide as much as useful metadata information which will enhance effective implementation.

Note:

Please don't maintain the spreadsheet separately from the data model, keep all your "Data Source"," Transformation Comments" and "Data Movement Rules" in the data model itself and generate the report as and when changes.