MANAGEMENT OVERVIEW

A GUIDE TO THE BENEFITS OF USING DATA MODEL EXTRACTION

Data Model Extraction with X-Analysis

Databorough Ltd.
66 York Road • Weybridge • UK
info@databorough.com
Phone +44 (0)1932 848564 • Fax +44 (0)1932 859211

A complete data model accurately describing all possible relationships between each file is essential for productive maintenance and development work. Such a model also provides the foundation upon which critical data administration tasks such as referential integrity testing, and test data extraction can be automated.

The Modeling tool set which extends to the basic X-Analysis performs automated data model extraction, analysis and documentation of (foreign keys) relational data model from RPG, COBOL or CA:2E. The Modeling tool set also provides the exports into UML, DDL, XML meta-data repositories and tools such as MyEclipseBlue, Rational, Erwin and Together.

Benefits

The Modeling tool set offers following benefits:

- Automated generation of data model through reverse engineering of legacy application
- Verifies relationships using actual program code and data
- Works without source code if necessary
- Makes accurate data modeling for even the largest systems possible
- Fully exportable to UML, DDL and other popular modeling tools
- Produces a data encyclopedia for entire system including field usages

And many more...

Here is a brief description of some of the features of the Modeling tool set which provides additional benefits to analysts and programmer.

The relational model of an enterprise application is an extremely powerful piece of information and potentially valuable asset to the organization. For almost all RPG or COBOL applications running on System i, there is no explicit data model or schema defined.

By the term model, we are referring to the foreign key or relational model, not just the physical model of the database.

Deriving Legacy Data Model

The Modeling tool set accomplishes the task of extracting the Data Model by analyzing the data structures of the physical and logical files, but it then programmatically traces these through all programs that use them to verify the existence of any cross-file relationships or foreign keys. These derived relationships can also be verified by the product by performing an integrity check on the actual data. This ensures that the data of the dependent file makes a reference, to data records from the owning file. In this way, the automated modeling process can fully extract the data model from even the most complex legacy system.

The relational model or architecture of the database can be reused in a number of scenarios including:

- Understanding application architecture
- Data quality analysis - referential integrity testing
- Automated test data extraction, scrambling and aging
- Building BI applications or Data warehouses

The Modeling tool set has the unique capability of automatically deriving the explicit system data model from a legacy RPG or COBOL application.

Automated Generation of Data Model

The Modeling tool set extracts information about the logic and structure of a legacy System i applications, and builds a reliable data-model from the data itself, augmented by the application logic and override files. The data-model thus generated is used for:

Visualization

Data-model can be viewed and analyzed in X-Analysis

Export to other CASE Tools

The Data Model can be exported to UML, DDL and other CASE tools such as COOL:Biz.

Generating Screen Functions

The generation of Screen Functions is incorporated into the XDMODEL command.

The Databorough's data model describes an application database in terms of its structure and how the database is used by an application.

Illustration 1: Modeling at a glance

Data Dictionary Analysis

The data dictionary based on the extracted data model, contains detailed information for every field in each file in the application database. Much of this data is the standard meta data extracted for each file and stored in the XDD file - for instance field and column names, field size and field type. Thus the record meta data is readily available for use by other applications.

Apart from extracting the meta data when building the data model, Modeling tool set also does the following:

Field Name	Type	Description
VFLD	10A	Field Name
VFILE	10A	PF Name
FLDCLS	10A	Field Class
SEQGRD	5 2	Sequence grid
SEQFSC	5 2	Sequence flat screen
FTYP	1A	Field type
ECDE	1A	Edit code
FGNK	1A	Foreign Key
BNAW	1A	Blank not Allowed
CDFD	1A	Code Field
YNFD	1A	Y/N Field
DATP	1A	Date Type
DATF	10A	Date Format
VPGM	10A	Validation Pgm
VPRM	1A	Vldn. Parameters
VCHK	1A	Chk blank entry
XVAL	1A	Further Validation
WPTYP	2A	Word Processing Typ
GATTR	1A	Global Attribute
WHCHD1	20O	Column heading 1
WHCHD2	20O	Column heading 2
WHCHD3	20O	Column heading 3
FLDHDG	50O	Field Heading
FDATTR	1A	Field Data Type
FLDLEN	3A	Field Length
FLDDEC	2A	Field Dec.Places
FLDBFP	5A	Input Buffer Posns.

Illustration 2: XDD the Meta Data file

Determine the format of the date in non-timestamps fields

Modeling tool set examines each field for date information. Indications of date information are the field name and the associated text with the field. This is re-enforced by examining the data to see if it contains date information and the format in which the date is stored, whether the date held is in year, month day order or month, day, year order and the number of digits used. This information is stored with the meta data.

Discover a field that will act as a descriptor for the record

A descriptor sums up the information on a record. This information is used when building the screen functions. Modeling Tool set can automatically determine which field is the most likely contender, using the field name, field type, field length and the associated text as clues.

Reverse engineer column and field sequence from application screens

This feature is available if the X-Analysis repository has been built. The fields and columns used in the screen displays are ranked in order of appearance and position in the screen displays. Those, which appear more often, are given higher ranking than fields, which are seldom used if at all.

View Data Dictionary

The Data Model once created by the Modeling tool set, may be browsed using the X-Analysis. All physical

files for the selected application would be listed along with their structure details. The various File-Field Details, Access Paths, File Relationships and Overrides information may be extracted.

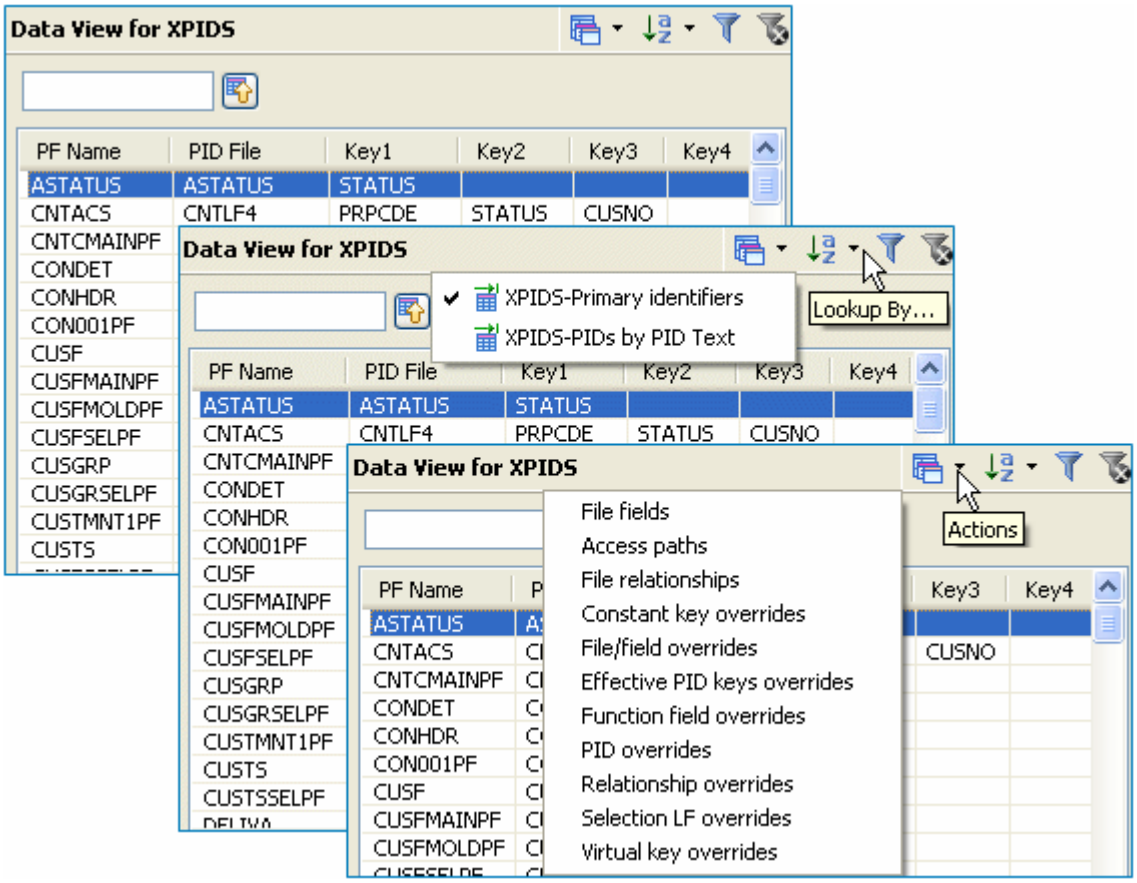

Illustration 3: View Data Dictionary

Entity Relationship Data Model Diagrams

The Data Model Diagram is the heart of the data modeling process and is done when all the primary identifiers have been found and the data dictionary has been built.

The Data Model Diagram displays the file relationships for a individual file or application/application area. The related members are displayed distinctly in yellow.

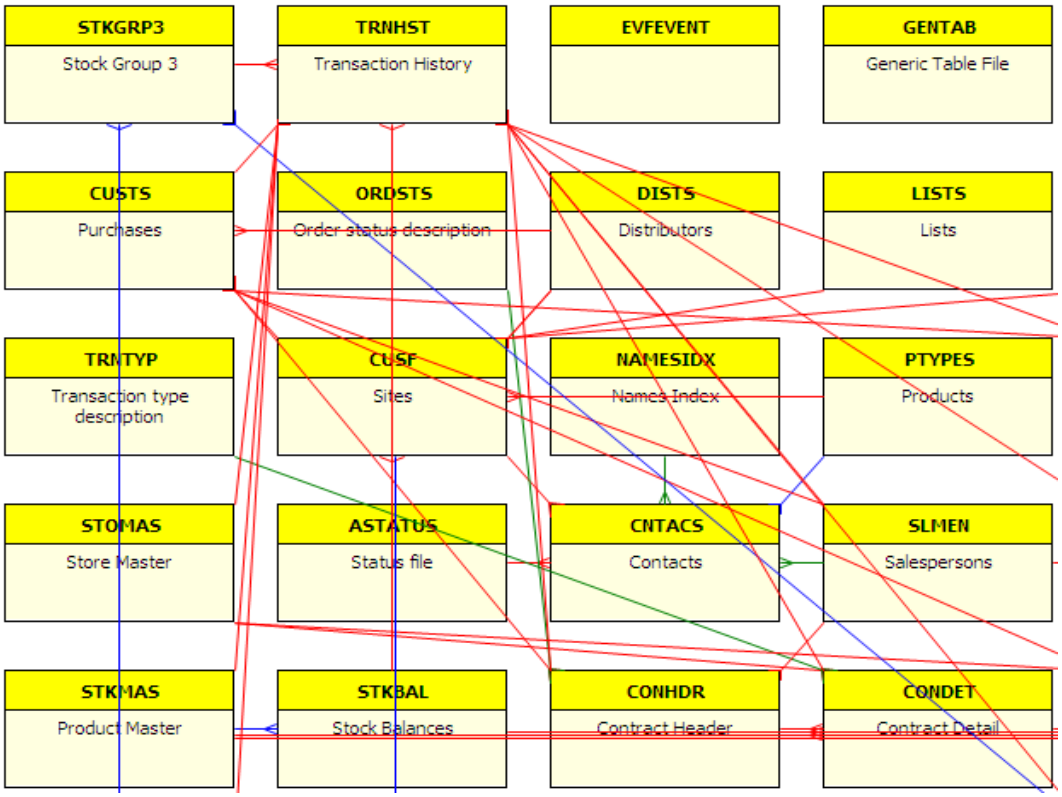

Illustration 4: Data Model Diagram

The Data Model Diagram also details the relations as displayed below.

DMD Relations for *ALL, Total Relations: 50

Rel No.	Dependent File	Relation Type	Parent File	Dependent Fields	Parent Fields	Dep LF
1	CNTACS	REFERS TO	ASTATUS	STATUS	STATUS	CNTLF3
2	CNTACS	REFERS TO	CUSF	CUSNO	CUSNO	CNTACS
3	CNTACS	REFERS TO	NAMESIDX	USERNM	IXNAME	
4	CNTACS	OWNED BY	PTYPES	PRPCDE	PRPCDE	CNTLF4
5	CNTACS	REFERS TO	SLMEN	SINIT	PERSON	
6	CONDET	REFERS TO	CONHDR	XWORDN	XWORDN	CONDET
7	CONDET	REFERS TO	STKBAL	XWABCD,XWAACS	XWABCD,XWAACS	CONDETL2
8	CONDET	REFERS TO	STKMAS	XWABCD	XWABCD	CONDETL3
9	CONDET	REFERS TO	STOMAS	XWAACS	XWAACS	CONDETL1
10	CONDET	REFERS TO	TRNTYP	XWRICD	XWRICD	
11	CONHDR	REFERS TO	CUSTS	XWBCCD	XWBCCD	CONHDRL4
12	CONHDR	REFERS TO	ORDSTS	XWSTAT	XWSTAT	
13	CONHDR	REFERS TO	SLMEN	PERSON	PERSON	CONHDRL2
14	CUSF	REFERS TO	ASTATUS	STATUS	STATUS	CUSFL2
15	CUSF	REFERS TO	DISTS	PRPCDE	DSDCDE	CUSFL6

Illustration 5: DMD Relations

The Data Model Diagram also displays the external objects (those objects which belongs to another application area). The external objects are displayed in blue.

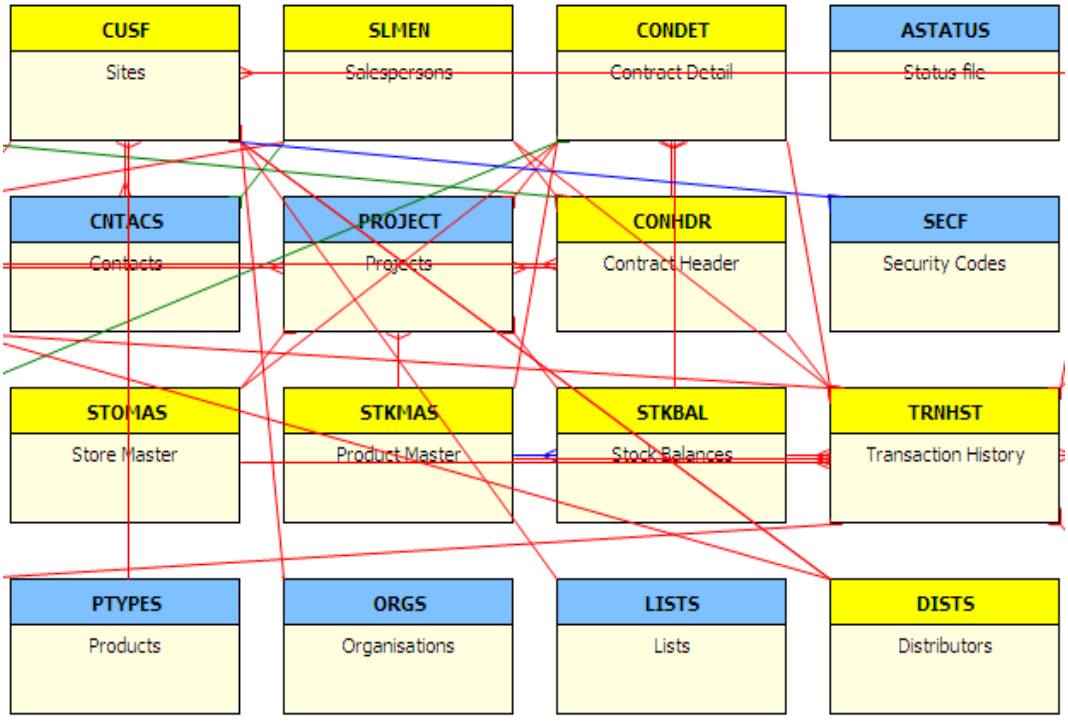

Illustration 6: Data Model Diagram with external relations

Types of file relationship

Owns
This is a PID-to-PID relationship - i.e. the PID of the owning file forms the first part of the PID of the dependent file.

Accesses
Here the PID of the owning file is contained in an access path other than the PID.

Refers to
This is the weakest of the relationships and is dependent by discovering the foreign keys in the dependent file. If foreign keys are required, every field is tested to see if it can be used as a key into another file by its PID. These relationships are flagged as reference only as it is not possible to access the dependent file from the owning file.

Once the PID has been identified, the other file relationships can be determined. The three types of relationship refers to, accesses and owns can be dealt with separately.

Data Browsing

Another important feature of the Modeling tool set is Data Browsing. File data may be browsed using X-Analysis. X-Analysis provides an 'View Data' option which displays the records of *FILE type objects (PF).

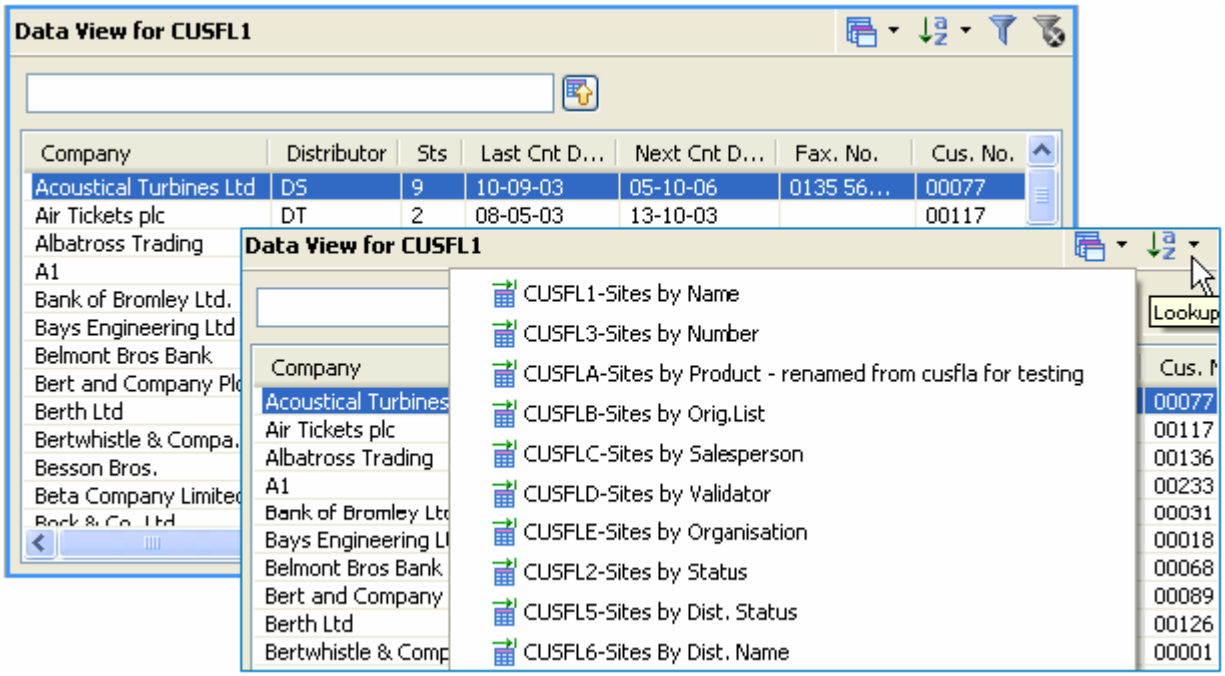

Illustration 7: Browse Data

The Modeling tool set process interrogates a host application and database and outputs the Data Model files which describe the internal structure and composition of the system. This information is then converted into a format more suitable for rapid application access.

Highlights

- Includes all features from X-Analysis Professional
- Automated generation of data model
- Browsing and analysis of data dictionary
- Entity relationship diagrams
- Data encyclopedia
- Builds Interactive, ready-to-use, drill-down application over legacy system
- Instant & Automated Structured, drill-down browsing of test/live data during testing and problem analysis
- Fully exportable to UML, DDL and other popular modelling tools
- Visio exports for graphical diagrams

The Modeling tool set extracts the Data Model information accurately which describes all possible relationships between each file. This is essential for productive maintenance and development work. The extracted data model also provides the foundation upon which critical data administration tasks can be automated.

Experience the fully loaded X-Analysis with 30 days trial copy of the software.
For any information regarding the X-Analysis please visit our web site:

www.databorough.com
or write e-mail to us at:
info@databorough.com

Databorough

© copyright Databorough 2010

Corporate Headquarters >

Databorough Ltd.
Weybridge Business Centre,
66 York Road,
Weybridge,
KT 129DY
United Kingdom
044-1932-848564
044-1932-859211
info@databorough.com
www.databorough.com

International Office >

Databorough Services
Suit# / Box# 504,
92 Caplan Avenue,
Barrie,
Ontario,
L4N 9J2
Canada
01705-458-8672
1800-605-5023 Toll Free
info@databorough.com
www.databorough.com