

B09 Managing Metadata in the Enterprise Using Information Catalog Manager
Roger D. Roles, Information Catalog Manager Architect, IBM

This talk will show how Information Catalog Manager (ICM), a component of DB2 Warehouse Manager, can be used to manage your enterprise metadata, both in conjunction with other DB2 tools such as Data Warehouse Center and DB2 OLAP Server, and with third party tools. We will demonstrate how Data Warehouse Center metadata can be published to the ICM, maintained using Information Catalog Center, and how to use extractors to populate your metadata catalog from relational databases which support JDBC, ERwin, ETI, Data Access Tools and other formats. We will demonstrate how to use the security features of the ICM which will allow you to partition your metadata catalog such that only authorized users will have access to sensitive information, and how to use the Information Catalog Center for the Web interface to allow end users to find data using common business terms with which they are familiar, understand what the data means, where it came from, and if it is up to date.

Session B09

Managing Metadata in the Enterprise using Information Catalog Manager

Roger D. Roles

rroles@us.ibm.com

A decorative graphic consisting of several green circles of varying sizes, some overlapping, arranged in a horizontal line. A central green rounded rectangle with a purple border is superimposed over the middle of this graphic.

IBM Data Management Technical Conference

Anaheim, CA

Sept 9 - 13, 2002

Abbreviations, Disclaimers, Trademarks

API - Application Programming Interface

DWC - Data Warehouse Center

ICC - Information Catalog Center

ICM - Information Catalog Manager

WHM - Warehouse Manager

All trademarks are owned by their respective trademark holders.



What is Information Catalog Manager?

- An application for tracking, managing, and using metadata (data which describes data)
- A component of DB2 Warehouse Manager
- A "Center" graphical user interface (UI design consistent with Control Center, Warehouse Center, etc.)
- A web graphical user interface for browsing, searching, and commenting metadata
- Default metadata object types and a sample information catalog

Sources of metadata

- Extraction, transformation, and loading tools use metadata to describe the processes (sources, transformations, destinations) used to build information warehouses
- Database engines use metadata to describe the information stored in the databases
- Data access and reporting tools use metadata to deliver information to end users
- OLAP tools use metadata to build multidimensional stores
- Case tools use metadata to model business processes

Using an information catalog

- Customize
 - ▶ Create/modify object types
 - ▶ Create relationship types
- Populate
 - ▶ Metadata publish to ICM
 - ▶ Metadata extraction
 - ▶ Metadata import
 - ▶ Interactive entry
- Use
 - ▶ Annotate metadata
 - ▶ Browse metadata
 - ▶ Navigate through related metadata
 - impact analysis
 - ▶ Search metadata and find data
 - ▶ Initiate applications to view data described by metadata

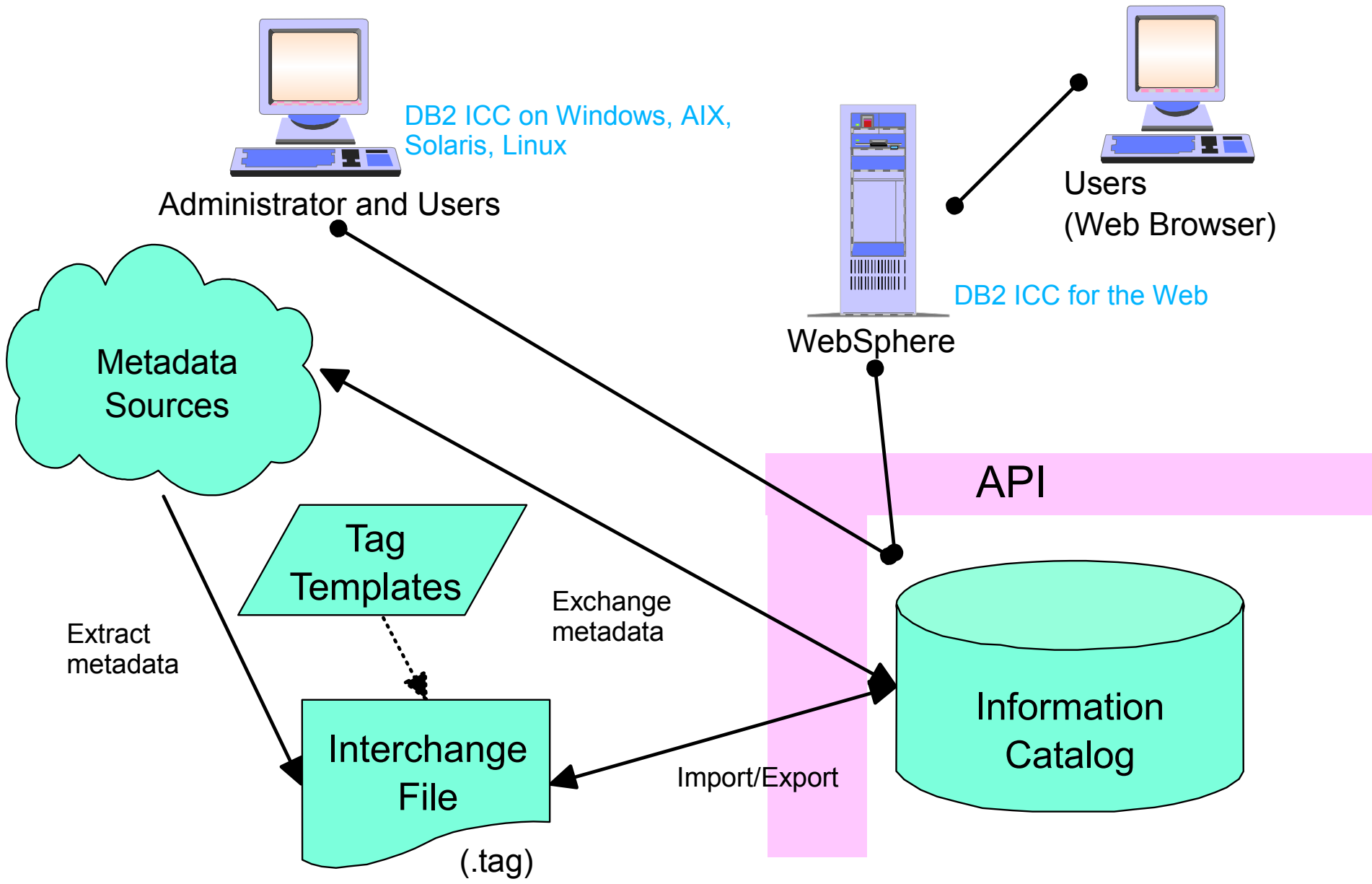
Questions answerable by an information catalog

- What data exists?
- Where is it located?
- What does it mean?
- Where does it come from?
 - ▶ How has it been transformed?
 - ▶ To what other data is it related, and how?
- When was the data last updated?
- Who is responsible for it?
- What format is the data in?
- What tools are available to retrieve and/or view the data?

End User Value

- Business users
 - ▶ Find data in the warehouse
 - ▶ Know if the data is current
 - ▶ Understand what it means
 - ▶ Define in business terms
 - ▶ Use applications to analyze it
 - ▶ Accessible from browser using web UI
- Warehouse builders
 - ▶ Keep track of data in the warehouse
 - ▶ Impact analysis
 - impact on warehouse of source changes
 - impact on users of warehouse changes

ICM Architecture



ICM Metadata Sources

- Data Warehouse publish
 - ▶ DWC to ICM
 - ▶ OLAP to ICM
- Relational database extractor (JDBC)
- Windows desktop applications
- IMS
- ERWin 3.5
- ETI
- Business partner applications
 - ▶ as provided by ISV
 - ▶ custom written
- Customer written applications

ICM V8 Changes

- Control Center style GUI on Windows, UNIX
- User preferences stored in IC
- Multiple concurrent administrators
- New search function
- Security by metadata object
- Custom relationship types
- Multiple property data types with constraints
- New predefined object types
- Easier and customizable Start Program
- Java API

ICM V8 Packaging

- The ICC GUI is installed with DB2 ESE admin tools
- The Manage Information Catalog Wizard is installed with DB2 Warehouse Manager
- Means - can install ICC with DB2 ESE, but to create the information catalog must have a DB2 Warehouse Manager license

Creating the information catalog

- Can be created during installation or after
 - ▶ Created during installation as a tools catalog (shared with Task Center) or separate catalog
 - ▶ Create catalog after installation using Manage Information Catalog Wizard (db2iccwz)
 - ▶ Wizard can also be used to migrate the data in a V7 catalog into a V8 catalog and to create legacy views
- Required information
 - ▶ Tools catalog name or database and schema
 - ▶ User ID and password for DB access
 - ▶ Option selections (groups, default types, TS)

Initialize - catalog selection

Manage Information Catalog Wizard

1. Task
2. **New Location**
3. Options
4. Summary

Specify the location of the new information catalog metadata.

Specify whether the new information catalog metadata will be in a DB2 tools catalog or another database.

To use a DB2 tools catalog, select the name of the system of the tools catalog that you want to use. The location of the database that contains the tools catalog for that system will be displayed.

You may also type names into the Database Name and Database Schema fields.

In a DB2 tools catalog In another DB2 database

Database name: icmdb **Locally cataloged name of DB to hold catalog**

Database schema: ICM3 **Schema name for catalog (later called catalog name)**

Database user ID: rroles

Database password: *****

◀ Back Next ▶ Finish Cancel

Initialize - options

Manage Information Catalog Wizard

1. Task
2. New Location
3. Options
4. Summary

Specify the options for creating a new information catalog.

You selected to prepare a new information catalog. Specify the default user group and the default power user group for the information catalog.

Default user group **Identifies users with browse authority**

Default power user group **Identifies users with the authority to create objects**

Select whether the Information Catalog Center's default object types should be created in this information catalog.

Create default object types

Specify the name of the table space to be used for the information catalog.

Table space

Initialization - command line

- Initialization of an information catalog may also be done using a command line utility - db2icminit
- More options available from the command line, but more complex to use
- May also be used to check the structure of the catalog, or to completely delete it
- Cannot initialize the user and power user groups - must set using ICC after creation
- Example: `db2icminit -db icmdb -s ICM3 -u rroles -p xxxx -api create -app skip def`

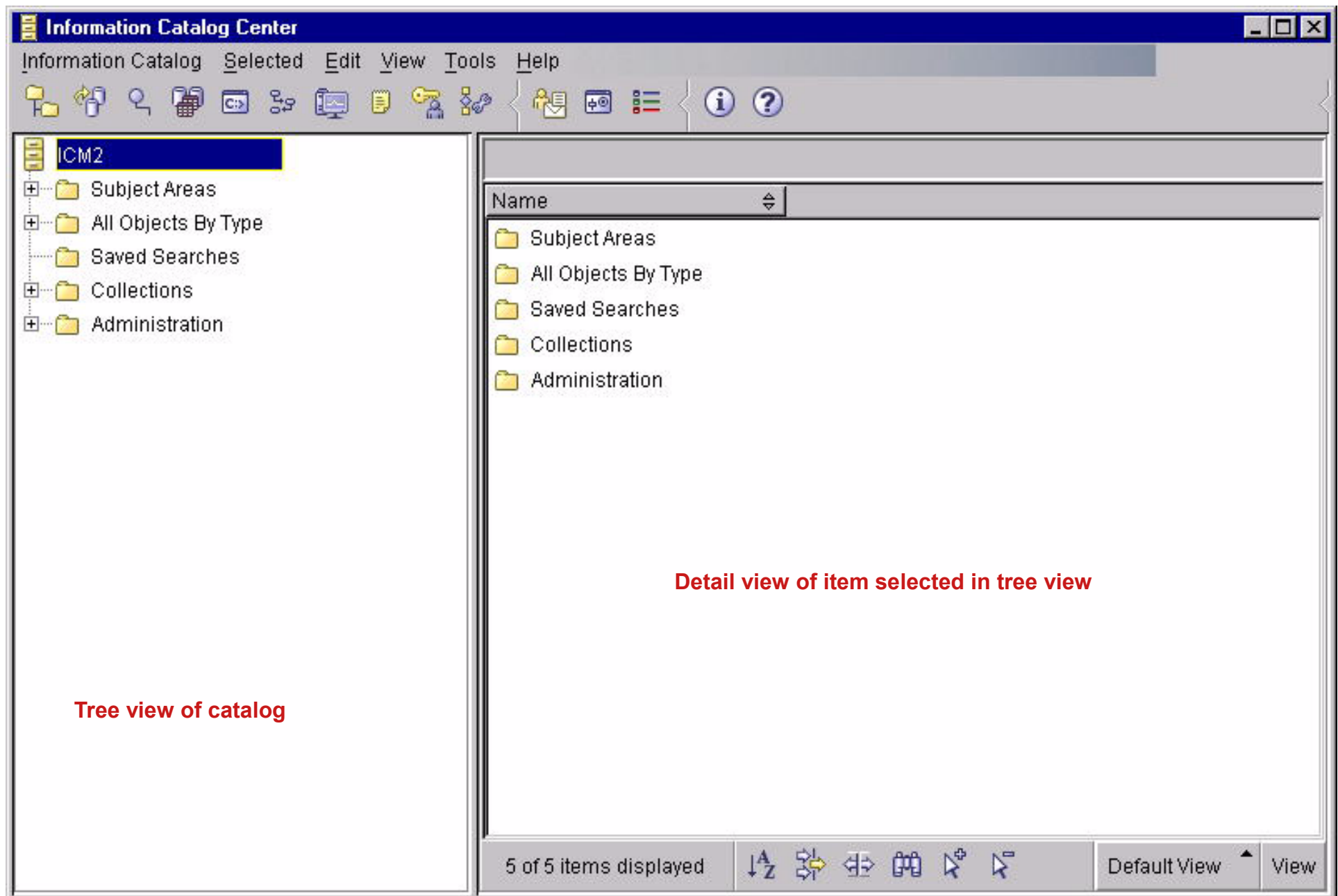
ICC User Functions

- Navigate a tree view of the catalog
- Search for objects with specific properties and save the search definitions for later reuse (stored in catalog)
- Find related objects by relationship (lineage)
- Add comments to objects
- Define collections of objects (stored in catalog)
- Initiate programs
- Set user preferences (stored in catalog)

ICC Admin Functions

- Define object types and create objects
 - ▶ Define properties
 - ▶ Set permitted values for properties
 - ▶ Set permitted relationship participation
 - ▶ Set default visible properties
- Define relationship types
- Control security
 - ▶ Set user and power user groups
 - ▶ Modify ACLs on objects
- Import/Export
- DB admins are ICM admins

ICC Main Window



Objects by Type Subtree

The screenshot displays the IBM Information Catalog Center interface. The left pane shows a tree view of object types, with 'All Objects By Type' expanded. The right pane shows a list of objects under the selected folder, including 'Subject Areas', 'All Objects By Type', 'Saved Searches', 'Collections', and 'Administration'. A red text annotation points to the 'Elements' folder in the left pane, stating 'Objects by type folder showing default object types.' Another red text annotation in the right pane states 'Detail still shows top level folders since the root (ICM) is still selected in the left panel.' The status bar at the bottom indicates '5 of 5 items displayed' and 'Default View'.

Information Catalog Center

Information Catalog Selected Edit View Tools Help

All Objects By Type

- Application data
- Attribute
- Audio clips
- Business subject areas
- Case Models
- Charts
- Column Mapping
- Columns or fields
- Comments
- Databases
- Dimensions within a multi-dimens
- Documents
- DWC Process
- Elements Objects by type folder showing default object types.
- Entity
- Files
- Glossary entries
- Images or graphics
- IMS database definitions (DBD)
- IMS program control block (PCB)
- IMS program specification (PCB)

Name

- Subject Areas
- All Objects By Type
- Saved Searches
- Collections
- Administration

5 of 5 items displayed

Default View View

Searches

- Search for date or text values
- Search for multiple text values
- Search all or a subset of object types
- Search name property or all properties
 - ▶ all text properties
 - ▶ all numeric properties
 - ▶ all text and numeric properties
- Match all (AND) or match any (OR) values
- Exact match or wildcard match
- Unlimited or limited maximum returned object count

Search Dialogs

Search - 2

From this window you can choose to search for two types of data: Text and Numeric or Date. Specify the appropriate search options. The more options that you specify, the narrower the search.

Search data type

Text and numeric Date

Type one or more search values separated by commas.

dave

Object types to search

- multidimensional databases
- OLAP Integration Server model
- Online news services
- Online publications
- People to contact**
- Presentations
- Programs that can be invoked from informa

Properties to search

Name only

All properties

Text properties

Numeric properties

Options

Specify how you want the search values matched.

Type of match

Match all values (AND)

Match any value (OR)

Logic of match

Match anywhere

Match exactly

Case sensitive

Maximum objects to return All

OK Cancel Help

Search - 2

From this window you can choose to search for two types of data: Text and Numeric or Date. Specify the appropriate search options. The more options that you specify, the narrower the search.

Search data type

Text and numeric Date

Select a date range to search

Between June 21, 2002 and June 21, 2002

Object types to search

- multidimensional databases
- OLAP Integration Server model
- Online news services
- Online publications
- People to contact**
- Presentations
- Programs that can be invoked from informa

Properties to search

Date created

Date last updated

All date properties

Maximum objects to return All

OK Cancel Help

Search Results Function

- Two panels in results dialog
 - ▶ Objects qualified by search in top panel
 - ▶ Parameters used for search in bottom panel
- Menu options to
 - ▶ Save the parameters of the search as a named search
 - ▶ Reopen the search dialog with the parameters and modify for a new search
- Saved search may be rerun to generate a new search results or opened so that parameters may be modified

Search Results

The screenshot shows a window titled "Search Results - 2" with a menu bar containing "Search results", "Selected", "View", and "Help". A dropdown menu is open over the "Search results" menu item, showing options: "Save...", "Search criteria...", "Page Setup...", "Print...", and "Close". Below the menu is a table with the following data:

Unique ID	Owner	Last Updated
Dave Salinero.Team leader of ...	RROLES	2002-06-17 15:06:40.384

Below the table, the text "Objects found matching search criteria." is displayed in red. At the bottom of the window, the search criteria are listed:

- Search data type =Text and numeric
- Search values =dave
- Object types =People to contact
- Properties to search =All text properties
- Logic to match =Match all values
- Type of match =Match anywhere
- Maximum objects to return =All

The search result "Dave Salinero.Team leader of product development" is shown at the bottom of the window. A red annotation "Criteria which produced this search result." points to the search criteria list.

Object Types

- The object type defines a collection of properties which characterize a type of metadata, and the roles that objects of the type may play in relationships
- Each property has
 - ▶ a name
 - ▶ a data type
 - ▶ optionally a list or a range of permitted values
- A subset of the properties may be defined as a "unique id" which may be used to distinguish objects of this type

Define Object Type

Define Object Type Wizard

1. Name and Id...
 2. Properties
 3. Unique ID
 4. Relationships
 5. Programs
 6. Summary

Specify the new object type's properties

Specify the new object type properties. In the Default properties table, specify which properties to show in other windows by selecting the Show check box. In the User-defined properties table, type the name, data type, and size of the property, and select whether the property is required and whether it is visible in other windows. You can click ... to see a list of permitted values that you can choose from.

Default properties

Property	Short Name	Data type	Length	Required	Show
Name	NAME	VARCHAR	200	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Owner	ICM\$OWNER	VARCHAR	30	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Creation user		VARCHAR	30	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Creation time		TIMESTAMP	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Last updated user		VARCHAR	30	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Last updated time		TIMESTAMP	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Creating application		BIGINT	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Application readers		BIGINT	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Application updaters		BIGINT	0	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

User-defined properties

Property	Short Name	Data t...	L...	Req...	Show	Value
Char Prop	CPROP	CHAR	10	<input type="checkbox"/>	<input checked="" type="checkbox"/>	aaa, bbb, ccc

Relationship Type

- Relationship types may be created in one of four categories, which govern the behavior of the relationship under modification
 - ▶ Hierarchical - roles parent and child
 - ▶ Peer to Peer - role object (link)
 - ▶ Precedence - roles preceding object, succeeding object (lineage)
 - ▶ Support - roles object and supported object (attach)
- Each relationship type has a set of object types which are permitted to take on each role

Define Relationship Type

Define Relationship Type [X]

Relationship type | Object type constraints

Select the object type and role associated with this relationship type

Available object types

Object Type	Role
People to contact	Child
Presentations	Parent
Presentations	Child
Programs that can ...	Parent
Programs that can ...	Child
Records	Parent
Records	Child
Relational tables a...	Parent
Spreadsheets	Parent
Spreadsheets	Child
Star Schemas	Parent
Star Schemas	Child

Selected object types

Object Type	Role
Databases	Parent
Relational tables and views	Child

OK Cancel Help

Access Authorization

- Each object created in the IC has an associated Access Control List which contains entries for each user or group with permission to access the object
- Each entry controls permissions to show, read, or write the properties of the object for one user or group
 - ▶ show - can see the existence of the object
 - ▶ read - can see the properties of the object
 - ▶ write - can change the properties of the object and modify the ACL
- Admins can see and modify everything




Object Privileges

Properties - CeIDial television commercial for 2000

C:\Program Files\SQLLIB\SAMPLES\SAMPDATA\example.avi.-

Properties Relationships Privileges

Select object privileges for each user or group. Click the Add... button to add more users or groups to the table. Note that any adjustments to the privileges of a database administrator will have no effect.

Users and Groups	Show	Show, Read	Show, Read, Write
 ICMUSER	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
 ICMUSER	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
 SHORTY	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Add...
Remove

OK Cancel Start Program... Help

ICC Web UI

- Traverse through object structure relationships using hyperlinks
- Starting point is Subject Areas
- Search for objects using name match or all properties match, text, numeric and date values, with wild cards
- Create comments on objects
- Invoke programs to render data via URL
- Similar to the V7 Web UI, but new implementation
- WebSphere Application Server application

ICC for the Web

Information Catalog Manager - Netscape

File Edit View Go Communicator Help

IBM Information Catalog Manager

Enter the search string. Search

Subjects
Advanced search
Home
Help

- ▶ [CelDial Products](#)
- ▶ [CelDial Sales Information](#)
- ▶ [Credit card accounts datamart database](#)
- ▶ [DWCGUI Subject 1](#)
- ▶ [Group of Lotus Approach objects](#)
- ▶ [INCOMTAR](#)
- ▶ [Local file warehouse target](#)

Description

CelDial Sales Information

Property	Value
Name	CelDial Sales Information
Short description	Objects dealing with sales of CelDial products
Long description	
Actions	-

Document: Done

Navigate and Results



ICM Sample Data

- A set of objects describing the metadata used at a fictitious company, the CelDial Corporation
- Used in the scenarios of the Information Catalog Manager Tutorial
- Requires a catalog created with the default ICM object types
- Is inserted into the catalog using the import function on the file
.../sqllib/samples/icmdemo/icmdemo.tag
- icmdemo.tag must be edited before import

ICM Relational Extractor

- Java program which extracts information from relational DBs and creates objects in the information catalog
- Creates database, table, and column objects with Contains relationships
- Can insert objects directly into the information catalog or write a tag format file
- Driven by a control file to select which information to extract
- Source provided for optional customization by user

Relational Extractor Example

- `.../sqllib/samples/icmjdbc`
 - ▶ README
 - ▶ `icmext.ctl` - sample control file
 - ▶ Java source
- Control file (`icm.ctl`)
 - ▶ `OUTFILE = ('m:\outfile.tag', REFRESH)`
 - ▶ `RDBNAME = (icmdb, roles, rroles)`
 - ▶ `DATABASE = (icmdb, rroles, xxxxx)`
 - ▶ `TABLE = (ICM, *)`
- `java JDBC2ICM -f icm.ctl -s ICM -c icmdb -u rroles -p xxxxx`

Relational Extractor Results

The screenshot displays the IBM Information Catalog Center interface. The left pane shows a tree view of databases, with 'icmdb' selected. The right pane shows a table of objects within 'icmdb', including names, owners, and last updated times. The status bar at the bottom indicates '23 of 23 items displayed'.

Name	Owner	Last Updated
APPLDATA00	RROLES	July 25, 2002 1:06:49 PM
APPLICATIONS	RROLES	July 25, 2002 1:08:29 PM
APPROACH00	RROLES	July 25, 2002 1:10:03 PM
ATTRIBUT00	RROLES	July 25, 2002 1:11:39 PM
AUDIO00	RROLES	July 25, 2002 1:13:14 PM
CHARTS00	RROLES	July 25, 2002 1:14:49 PM
COLUMN00	RROLES	July 25, 2002 1:16:23 PM
COLUMNMAP00	RROLES	July 25, 2002 1:18:02 PM
COMMENTS00	RROLES	July 25, 2002 1:19:42 PM
CONTACT00	RROLES	July 25, 2002 1:21:20 PM
DATABASE00	RROLES	July 25, 2002 1:22:58 PM
DGNEWS00	RROLES	July 25, 2002 1:24:36 PM
DIMENSION00	RROLES	July 25, 2002 1:26:14 PM
DOCS00	RROLES	July 25, 2002 1:27:51 PM
DWCPROC00	RROLES	July 25, 2002 1:29:25 PM
ELEMENT00	RROLES	July 25, 2002 1:30:59 PM
ENTITY00	RROLES	July 25, 2002 1:32:35 PM

DWC Publish to ICM

- Allows selected DWC information to be transferred to the ICM catalog
- Once information is transferred, DWC maintains a "Publish" object which can be used to update the information in the catalog, or schedule regular updates
- The ICC can then be used to annotate and browse the information about DWC processes
- Publishable information includes Subjects, Sources, Targets, Steps, etc.

DWC to ICM Object Mapping

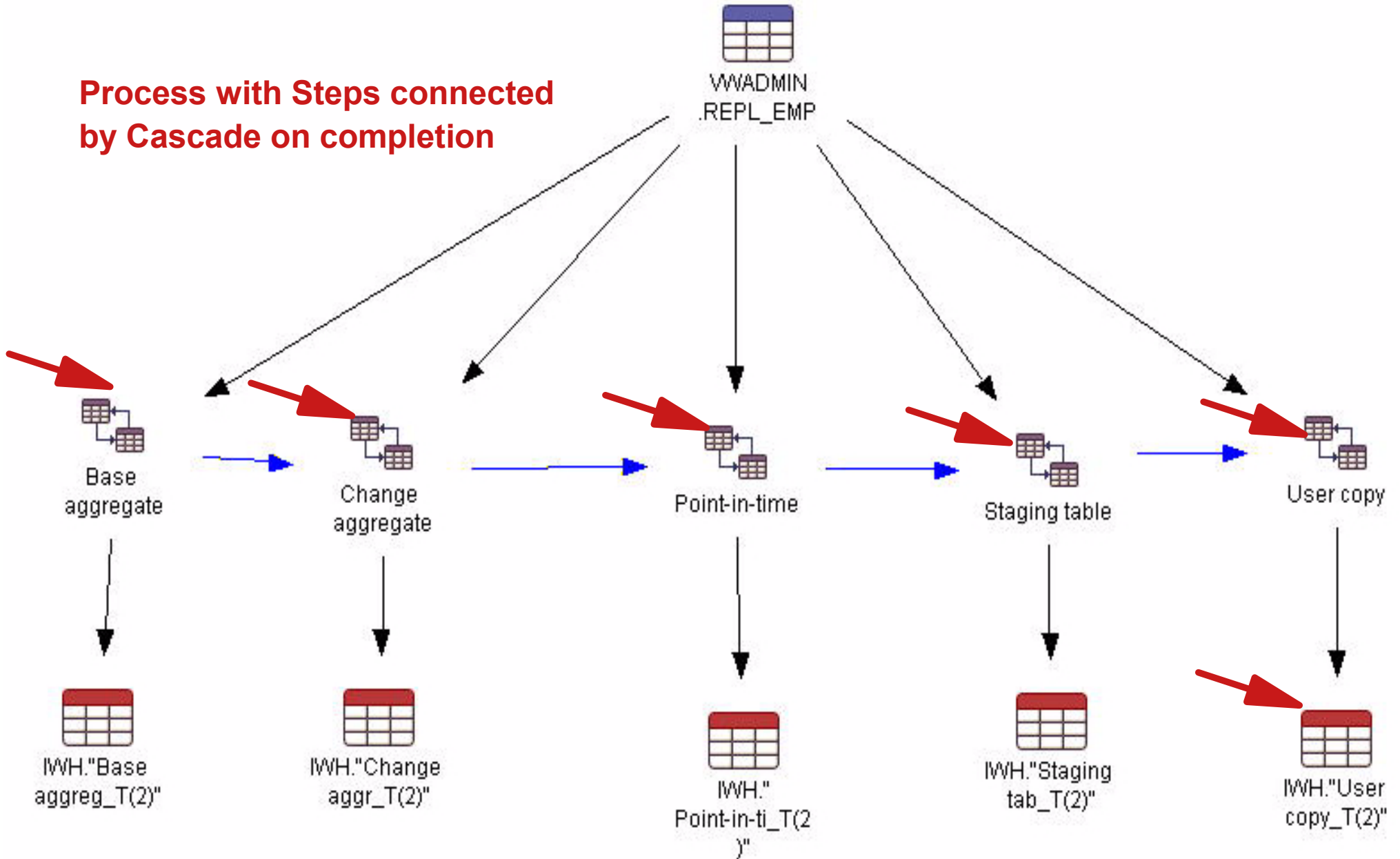
DWC	ICM
Column or field	Columns or fields
Map	Column mapping (new in V8)
Step	Transformation (at the table or column level)
Subject	Subject areas
Table, file, or segment	IMS Segments, relational tables, views
Warehouse schema	Star Schema
Warehouse source or Warehouse target	Database, IMS database definition

DWC Publish to ICM Example

- DWC process with Steps connected by Cascade on completion relationships
- Publish to column level to ICM
- Note the target of the last step - IWH."User copy_T(2)"
- We will examine this in ICM after publish

DWC Process Modeler

Process with Steps connected by Cascade on completion

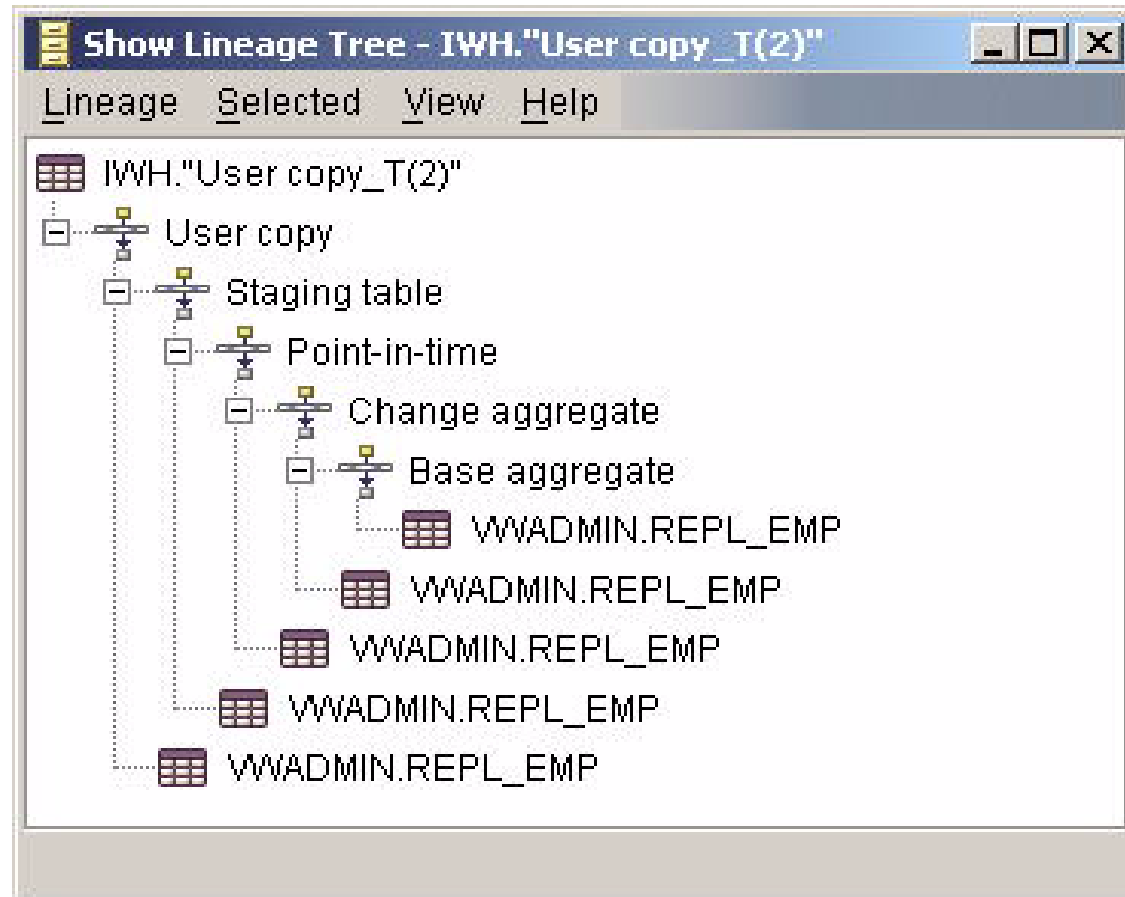


DWC Process in ICM

The screenshot displays the IBM ICM interface. On the left is a tree view of 'Subject Areas'. The selected area is 'DWC GUI WVP PANELS Subject', which is expanded to show 'Replication Process'. Under 'Replication Process', the table 'IWH."User copy_T(2)"' is highlighted with a yellow box. A red arrow points from this table to the right-hand pane, which displays a list of columns for the selected table. The columns are: BIRTHDATE, EDLEVEL, EMPNO, FIRSTNME, HIREDATE, JOB, LASTNAME, MIDINIT, PHONENO, SALARY, SEX, and WORKDEPT. A red arrow points from the text 'Result table columns' to this list. Another red arrow points from the text 'Result table' to the highlighted table in the tree view.

Name	Owner	Last Updated
BIRTHDATE	ROBINNT	July 16, 2002 3:34:07 PM
EDLEVEL	ROBINNT	July 16, 2002 3:34:07 PM
EMPNO	ROBINNT	July 16, 2002 3:34:07 PM
FIRSTNME	ROBINNT	July 16, 2002 3:34:07 PM
HIREDATE	ROBINNT	July 16, 2002 3:34:07 PM
JOB	ROBINNT	July 16, 2002 3:34:07 PM
LASTNAME	ROBINNT	July 16, 2002 3:34:07 PM
MIDINIT	ROBINNT	July 16, 2002 3:34:07 PM
PHONENO	ROBINNT	July 16, 2002 3:34:07 PM
SALARY	ROBINNT	July 16, 2002 3:34:07 PM
SEX	ROBINNT	July 16, 2002 3:34:07 PM
WORKDEPT	ROBINNT	July 16, 2002 3:34:07 PM

ICM Show Lineage Tree



Show Precedence Relationships

The screenshot shows a dialog box titled "Show Related - Staging table" with a close button (X) in the top right corner. Below the title bar is a dropdown menu showing "Transformations - Staging table". A warning message reads: "WARNING: Only first level dependencies are shown." Below the warning are four tabs: "Hierarchical", "Peer to peer", "Support", and "Precedence", with "Precedence" being the active tab. The main area contains a table with three columns: "Name", "Relationship Type", and "Role".

Name	Relationship Type	Role
User copy	Cascade	Succeeding obj
Point-in-time	Cascade	Preceding objec
WWADMIN.REPL_EMP	Input	Preceding objec
IWH."Staging tab_T(2)"	Output	Succeeding obj

At the bottom of the dialog box are "Close" and "Help" buttons.

Show Hierarchical Relationships

WARNING: Only first level dependencies are shown.

Hierarchical Peer to peer Support Precedence

Name	Relationship Type	Role
EMPNO -> EMPNO	Contains	Child
FIRSTNME -> FIRSTNME	Contains	Child
MIDINIT -> MIDINIT	Contains	Child
LASTNAME -> LASTNAME	Contains	Child
WORKDEPT -> WORKDEPT	Contains	Child
PHONENO -> PHONENO	Contains	Child
HIREDATE -> HIREDATE	Contains	Child
JOB -> JOB	Contains	Child
EDLEVEL -> EDLEVEL	Contains	Child
SEX -> SEX	Contains	Child
BIRTHDATE -> BIRTHDATE	Contains	Child
SALARY -> SALARY	Contains	Child
Replication Process	Contains	Parent

Close Help

Resources

- DB2 books
 - ▶ Information Catalog Center Administration Guide
 - ▶ Information Catalog Center Tutorial
 - ▶ DB2 Warehouse Manager Installation Guide
 - ▶ Data Warehouse Center Application Integration Guide
- URLs
 - ▶ <http://www.ibm.com/software/data/db2>
 - ▶ <http://www.ibm.com/software/data/db2/bi>
 - ▶ <http://www.ibm.com/software/data/db2/>
 - datawarehouse