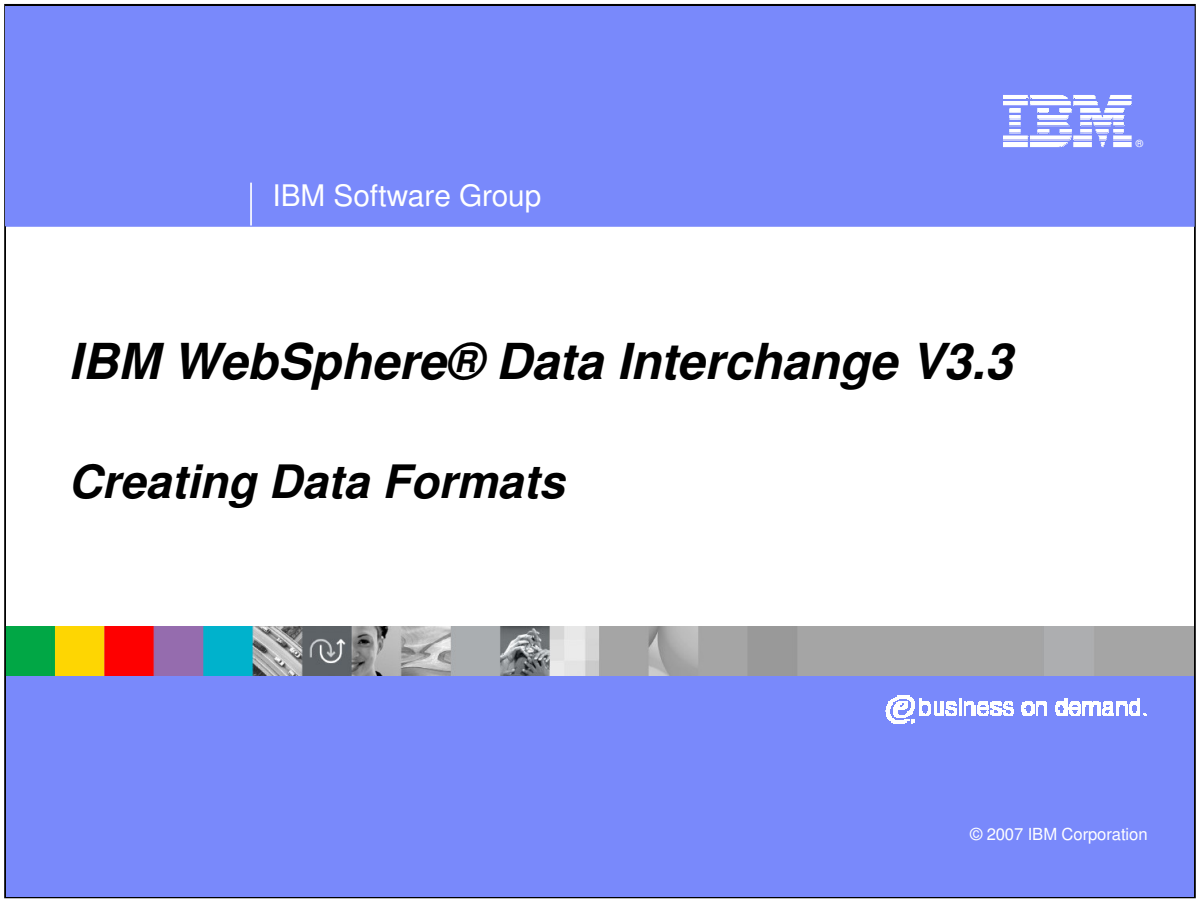


This presentation will demonstrate how to create Data Formats.

Agenda

- Review Data Formats
- Demonstrate how to create each Data Format Component.
- Review the Record Id Information Object.

The presentation will demonstrate how to create Data Format Components.

Creating Data Formats

- Defines the layout of your application data
- Records, field names, and lengths
- Repeating structures and loops

The term *data format* defines the layout of your application data. It is a document definition. The word *data* refers to the information itself. The word *format* refers to the physical layout of information in the file, such as field names and lengths. WebSphere Data Interchange requires a description of the data format for each business application that generates data for translation, or uses translated data. Application data must be described to WebSphere Data Interchange so that it can be used as either a source or target for translation.

Creating Data Formats

- Defining Data Formats
- Create Data Format Dictionary
  - Create Data Format Definitions or Document
    - Create Record definitions
      - Create Structures
        - Create Fields
      - Create Fields
    - Create Loops
  - Import Cobol Copybook
    - Imports records, structures, fields, creates code lists

The concepts for defining a data format are similar to Electronic Data Interchange (EDI) Standards. You Define a Data Format Dictionary. The Dictionary contains components for field, structure, record, and loop definitions. The data format definition contains record and loop definitions for the business document layout. Records can contain structures which contain fields and fields. All the components within a Data Format Dictionary can be copied, updated, and deleted and all components can be re-used in different business document definitions. For example a record can be used in 2 different data format definitions.

COBOL copybooks can be imported into a Data Format Dictionary. You can use this mechanism to create or update Data Format Record, structures, fields, and code lists. The imported Records, Structures, and Fields will be a part of the Data Format Dictionary into which they are imported. The Data Format Records can be used in a existing Data Format or a new Data Format.

Creating Data Formats

- Application Business Document Definitions
  - Defined as Raw data or C&D format
  - Records
  - Structures not a record (Occurs within a record)
  - Fields

You usually need to create a data format for every unique business document that is used or created by WebSphere Data Interchange (WDI). A single data format can be mapped to multiple documents. A data format can be defined as raw data or Control and Data (C&D) format. Raw data means there is a record identification that uniquely identifies each record and each record is fixed length. Raw data formats can also be delimited for example comma separated values. C&D format is a WDI defined format. The C or Control Record in the input data contains information for the data format to use for parsing the input, trading partner information, and override information. Each C record in the input data will signal a new message. Each D or Data Record in the input data begins with the value 'D' followed by the Record name used when the data format was defined followed by the data for the record. Records contain fields and structures. Fields are individual values and structures are grouped fields that may or may not repeat within a record.

Creating Data Formats

- C&D format example

```
CDIUSER              DIUSERWORKSHOP          YY N
DRECORD10      24100000PONUM                 19990818
DRECORD11      BY 9999
DRECORD11      SE V00000WWWW
DRECORD11      DP XXXX
DRECORD20      1    0000042XYZ               BP 92 XXXX
DRECORD20      2    00000XXXXXXX             BP 92 TEST
DRECORD21      4  000000000001000PCE19990906
DRECORD21      4  000000000002000PCE19990913
DRECORD21      4  000000000001000PCE19990906
DRECORD21      4  000000000002000PCE19990913
DRECORD21      4  000000000012000PCE20000501
DRECORD21      4  000000000011000PCE20000601
DRECORD21      4  000000000010000PCE20000701
```


This is an example of C&D input data. The Control or C record is pre-defined and is documented in the WebSphere Data Interchange Version 3.3 Utility Commands and File Formats Reference.

Creating Data Formats

- Raw Data Fixed length records example

```
H00 ROHIT       ZZ003020VICS  98052110220000017080000061920
H01 00121212121212              0007142024190002
SH011                  S
SH02CNT250001984                                                0000000000
SH03B               M ABCD LEASING INC              CC                      0000
SH06011940715
SH07STHLLOOP, INC.                  9 DUNSNUMBR0897
SH07SFTEST COMPANY PDQRST           9 0023233332222
SH08123419 STREETS AVENUE
SH09WALKER                  GA49503
OR012                  O
OR02896261                          940712
IT013                  I
IT02        UA011111005970              UP111110059700
IT03        0000000012CA00000000000000012CA
IT0400000100001200EA      000000000000000  000000000000000
IT014                  I
IT02
IT03        0000000022CA00000000000000022CA
IT0400000100001200EA      000000000000000  000000000000000
```


This is an example of Raw data fixed length records.

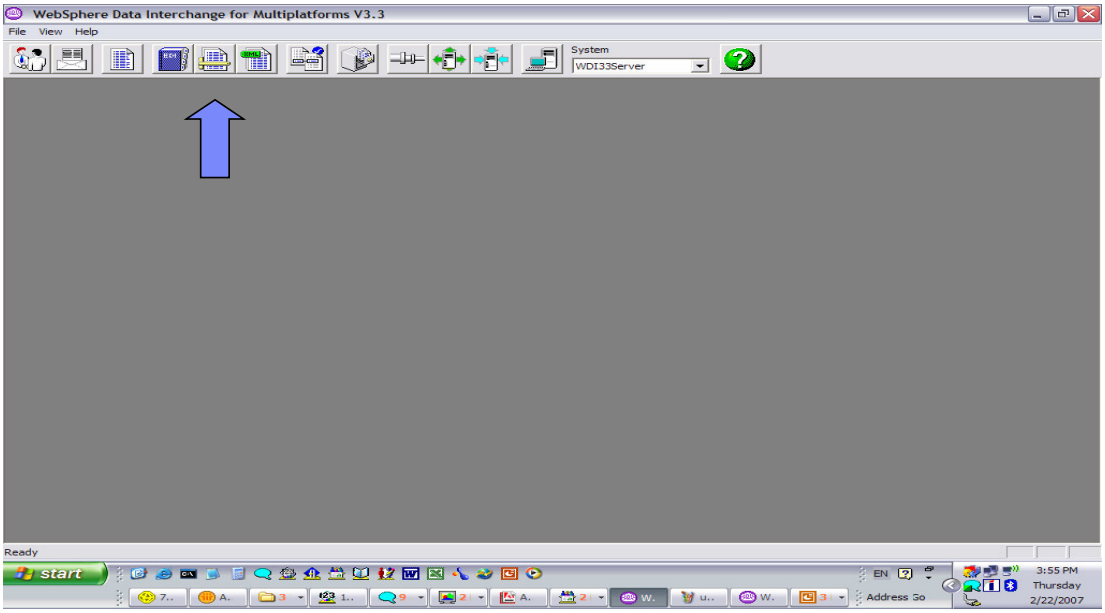
Creating Data Formats

- Raw Data delimited example

"(201) 207-0109","Record 1 / N/A",12/20/2004,"AC149",$0.00,120,$0.00,$149.99,$11.94
"(201) 207-4445","Record 2 / N/A",12/20/2004,"NE300",$0.00,470,$30.80,$35.00,$0.00

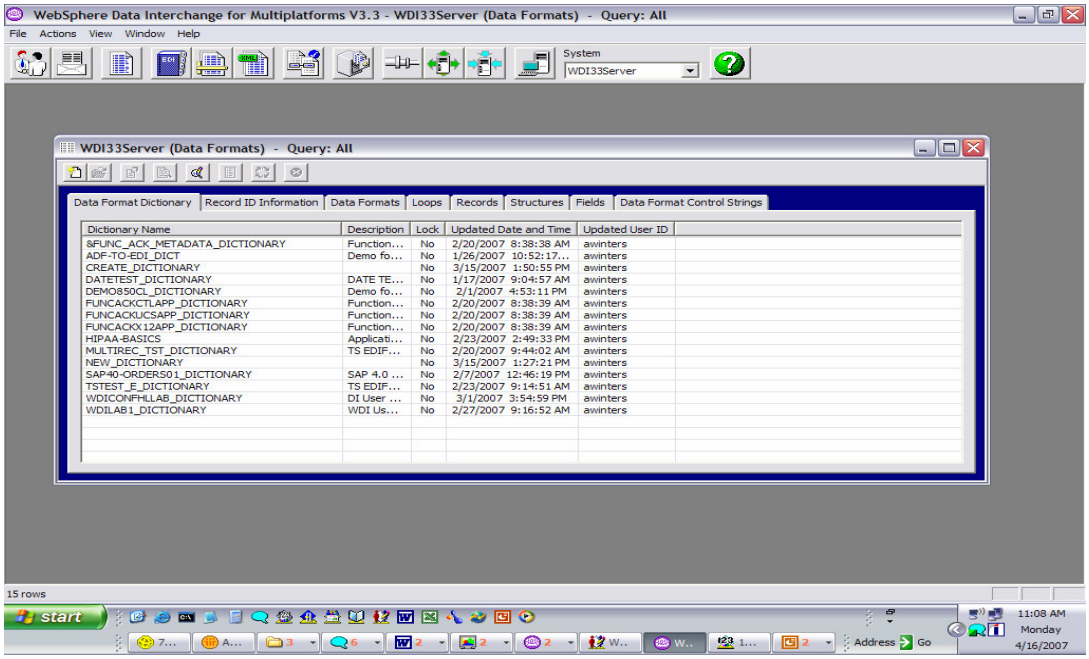
This is an example of Raw data comma delimited records.

Data Formats



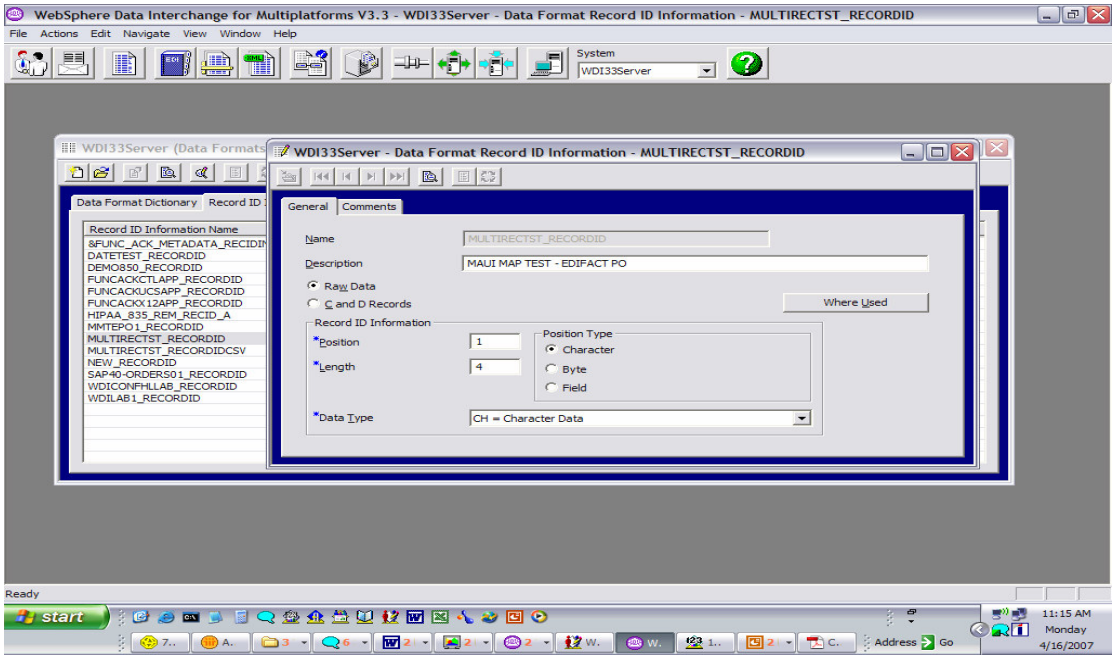
This is the location of the WebSphere Data Interchange (WDI) Client Data Formats Functional Area.

Data Formats



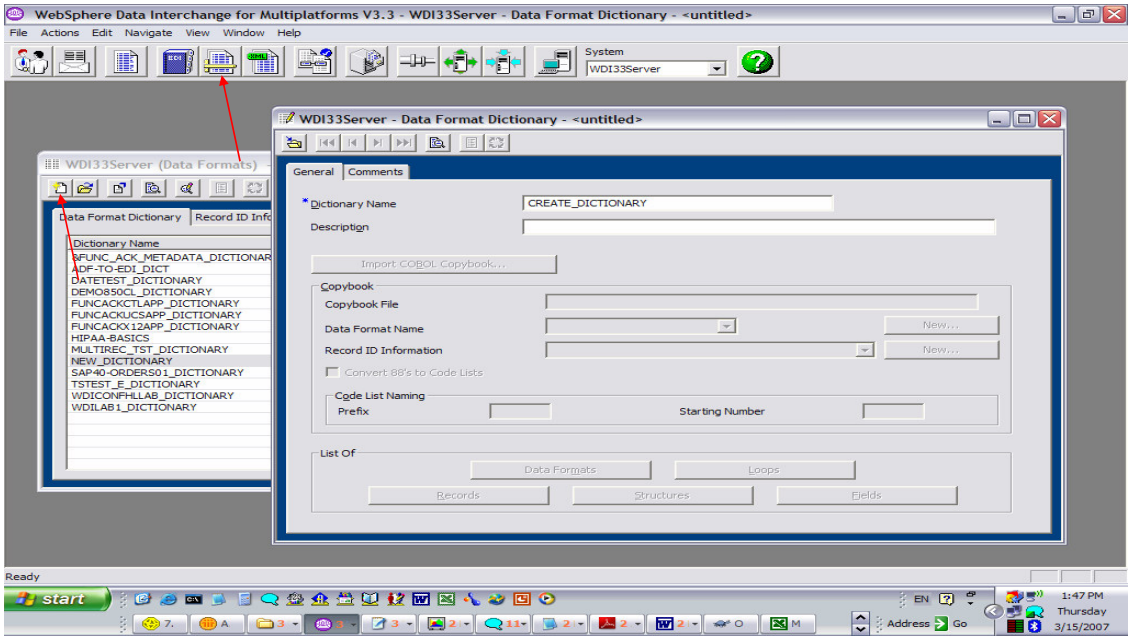
Components of Data formats are dictionary, record ID information, data formats, loops, records, structures, and fields. These are similar to Electronic Data Interchange (EDI) Standards components but describe your application data. The dictionary contains the component definitions for fields, structures, records, loops which allows you to re-use components within different data format definitions. Data format definitions contain records and loops. Loops contain records. Records contain fields and structures. And structures contain fields.

Data Formats



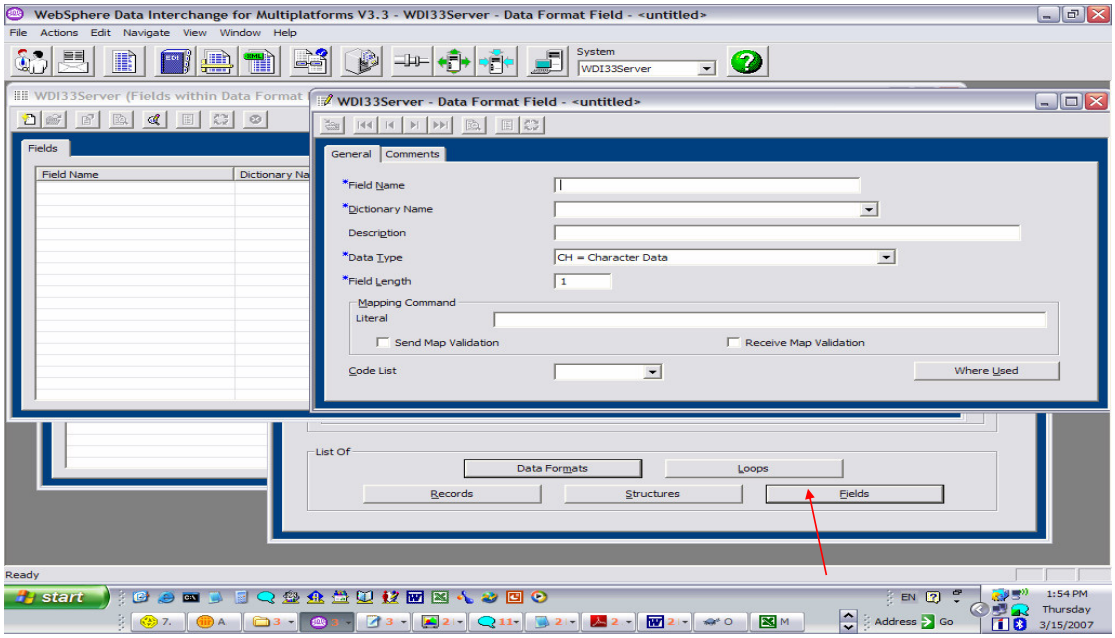
Raw data as opposed to C&D format is identified in the Record ID Information object.

Creating Data Formats



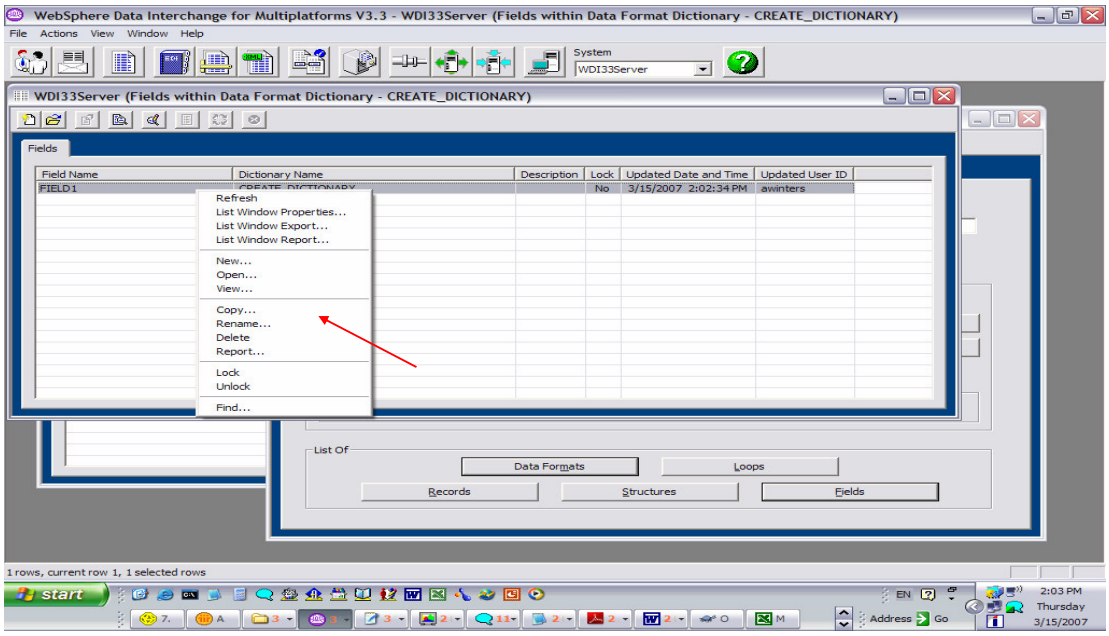
Data Formats are located in the Data Format functional area. The first step is to create a data format dictionary. Select the New button from the Data Format editor. Enter the Dictionary Name and save.

Creating Data Formats



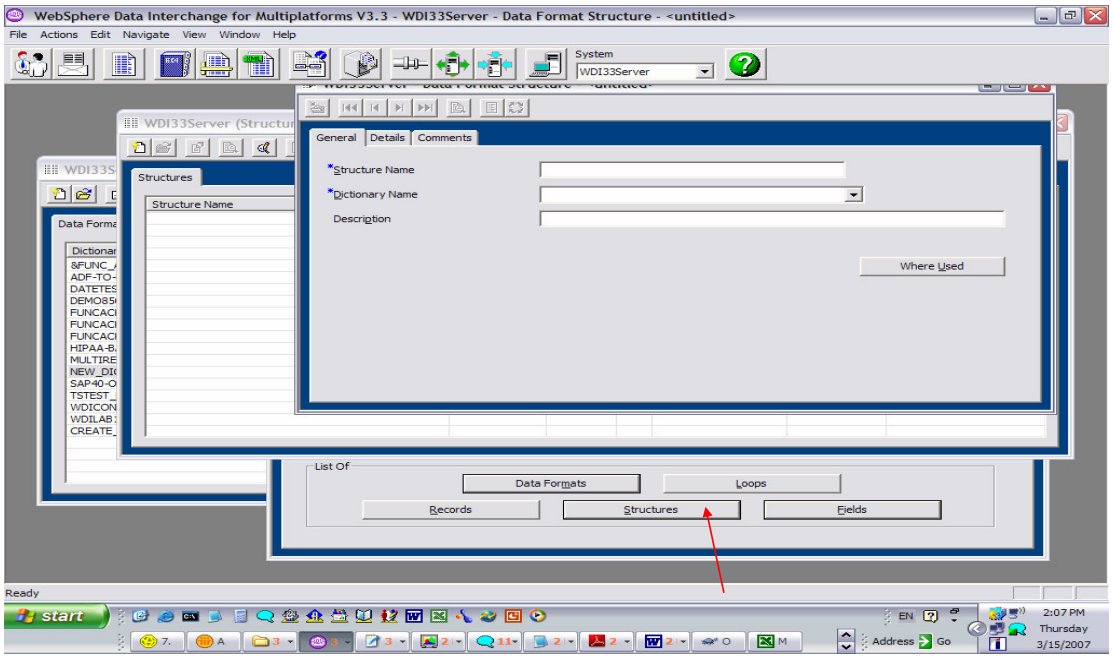
You begin with defining the fields. Fields are fundamental pieces of data, such as prices or item numbers or first names. Enter the name and select the Dictionary where this field will be defined. Select the data type and length for this field. The Mapping Command and code list specification are used for Send and Receive maps only. This allows you to associate a literal or mapping command or a code list with this field. When the field is used in a map the literal, command, or code list will be copied into the map.

Creating Data Formats



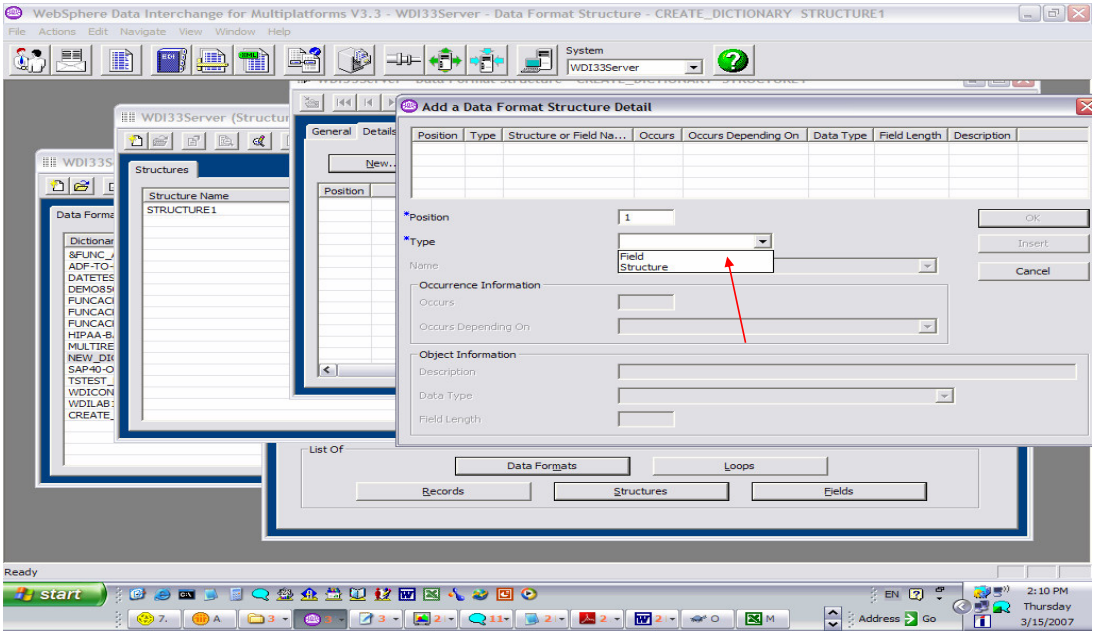
You can right click in the list window to select options. In addition to opening and viewing elements, you can copy, rename, and delete elements.

Creating Data Formats



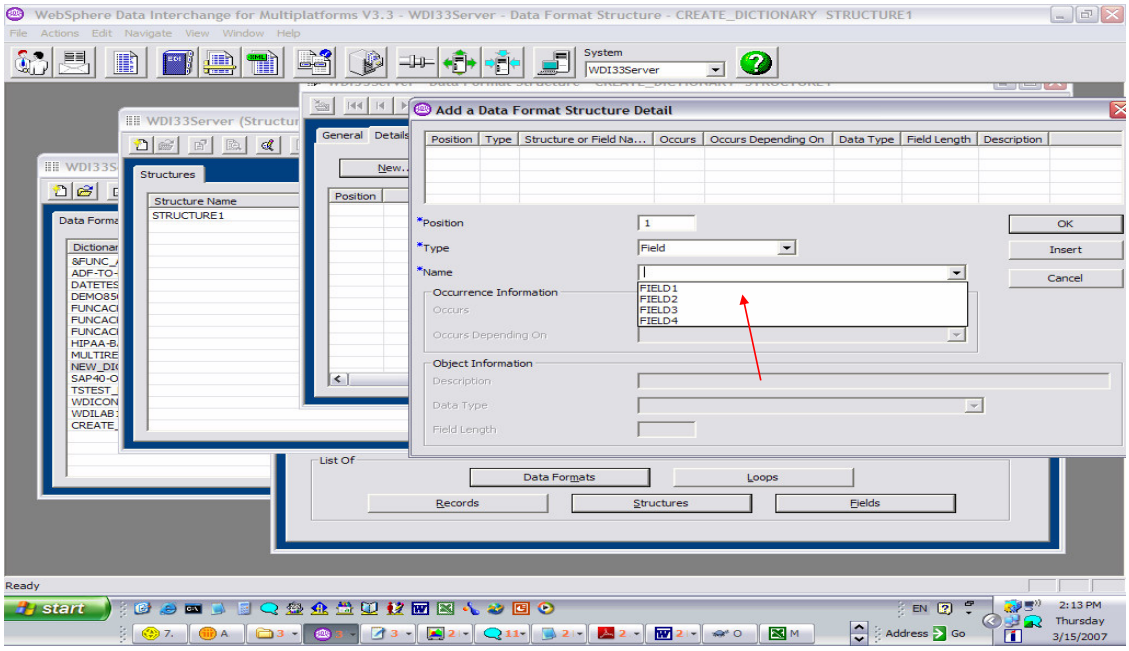
A Data Format Structure is a group of related Data Format Fields, which is probably unique to your company. When multiple Data Format Fields always appear together, you can designate the group as a Data Format Structure and give it a Data Format Structure name. Enter the Structure Name and select the Dictionary. Press the Save button and move to the detail tab to define the related fields.

Creating Data Formats



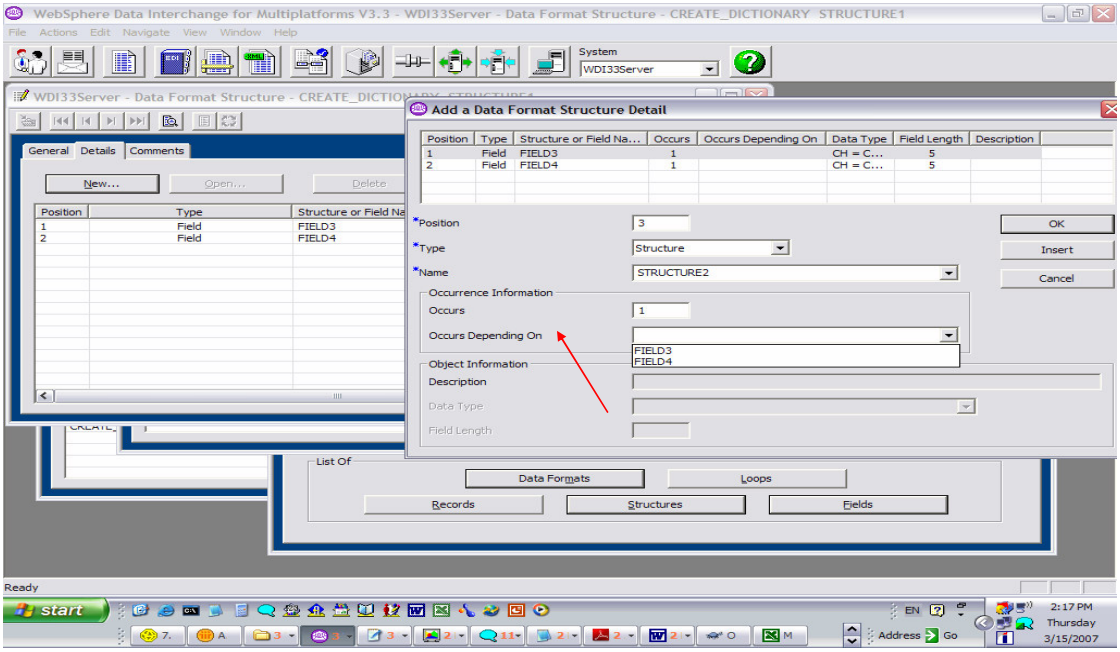
Data Format Structures can contain Data Format Fields and other Data Format Structures.

Creating Data Formats



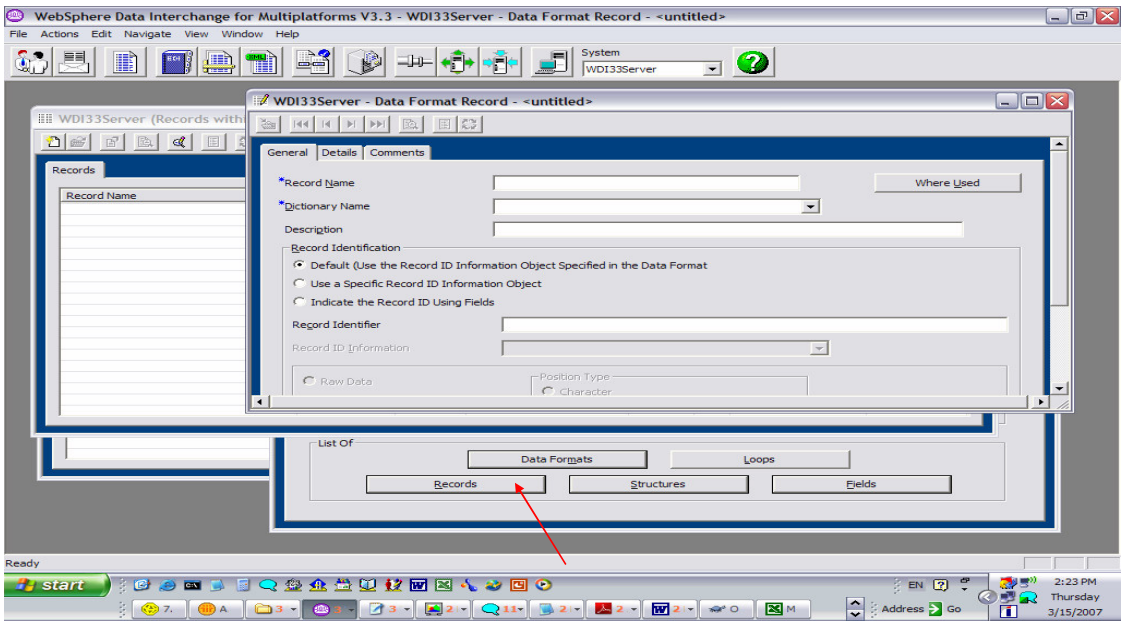
A list of fields and structures is selectable using the drop down list.

Creating Data Formats



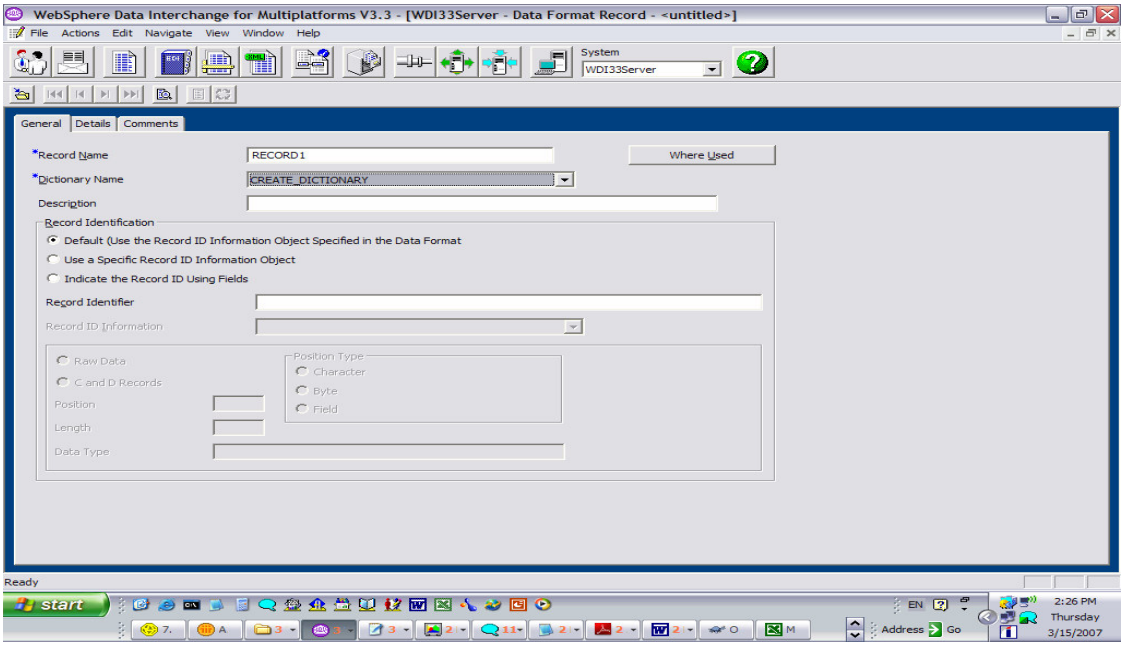
The Occurs field indicates the maximum number of times a Structure repeats consecutively. Enter a value from 1 to 32767. This column is disabled and ignored when the row represents a Field or when a value is in the Occurs Depending On column. The Occurs Depending On field indicates that a value within the document will determine how many repeats a Structure will have. To do this, select the name of an existing Data Format Field within the Structure that defines the field containing the value within the document.

Creating Data Formats



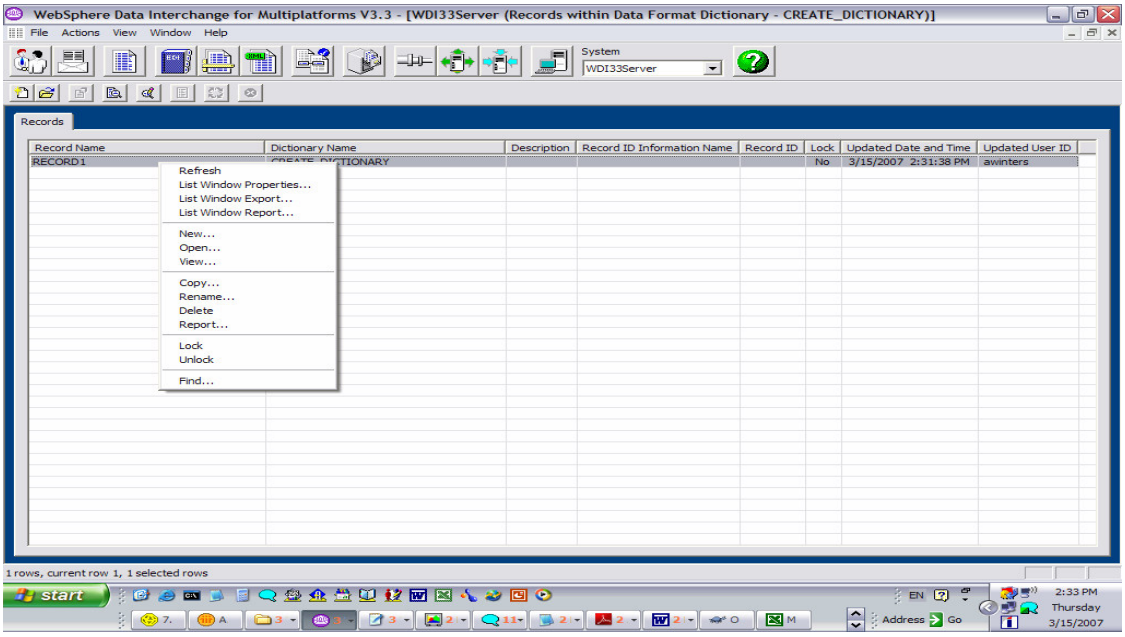
A record is a set of related Data Format Fields and structures as they are defined in an application's data. Every record in your application must be defined in your Data Format.

Creating Data Formats



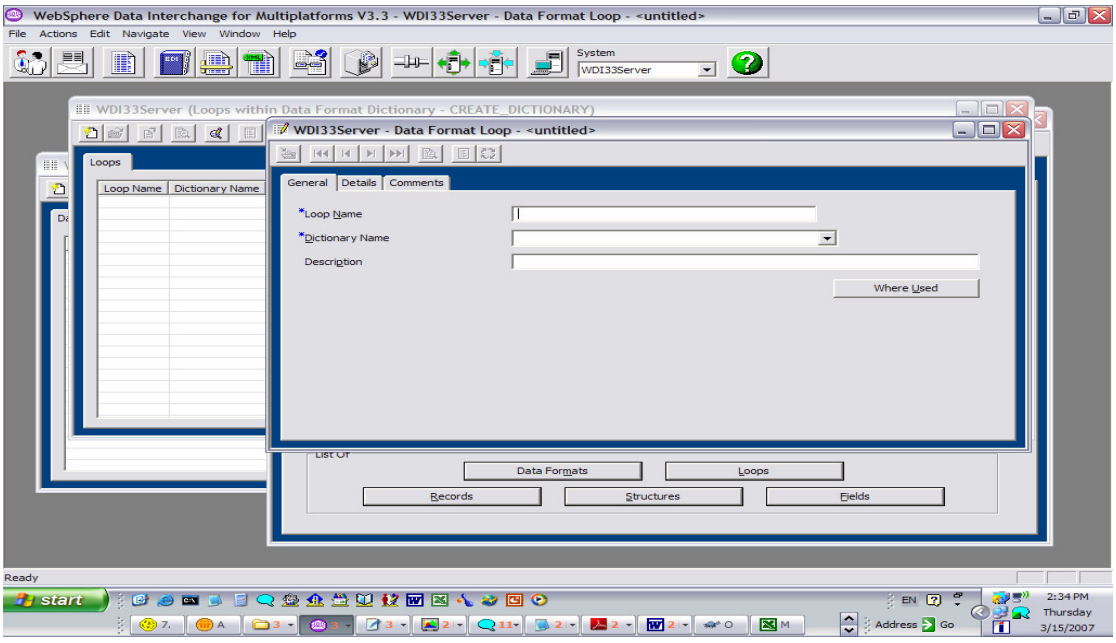
Enter the Record name and dictionary name. We will discuss the Record Identification later in this presentation.

Creating Data Formats



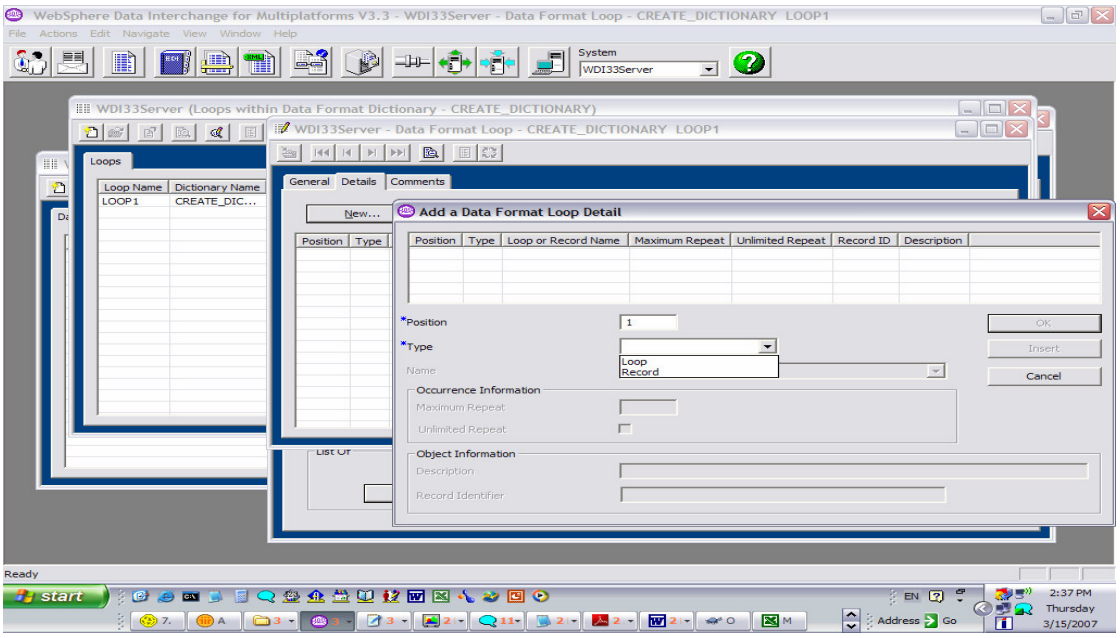
The same options are available on the list windows throughout the Data Format editors.

Creating Data Formats



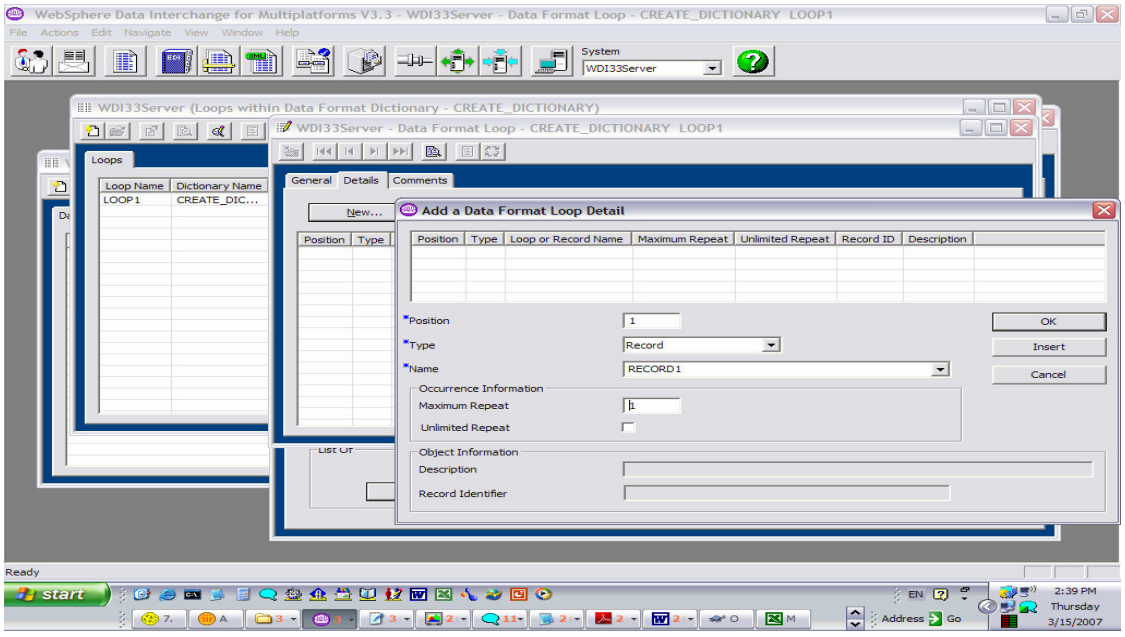
A loop is a group of records that repeat up to the number of times specified in the Maximum Repeat column in its editor. Loops can appear within other loops; this is referred to as a nested loop. Enter the Loop Name and dictionary name and press the Save button. Then move to the Details Tab.

Creating Data Formats



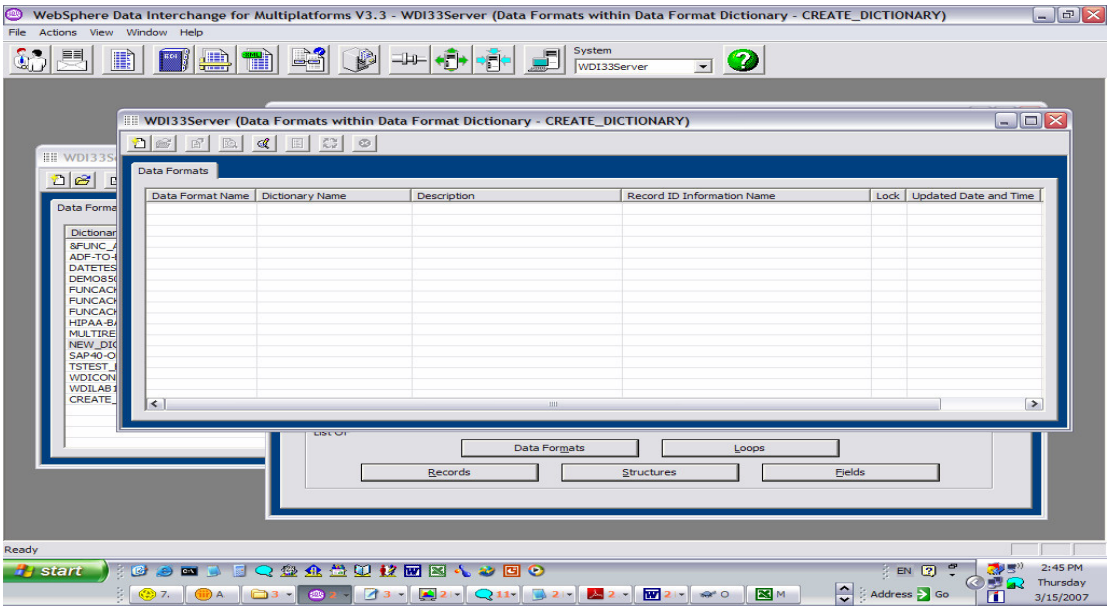
Use the Data Format Loop Editor to enter new Data Format Loops into a Data Format Dictionary or to edit existing Data Format Loops.

Creating Data Formats



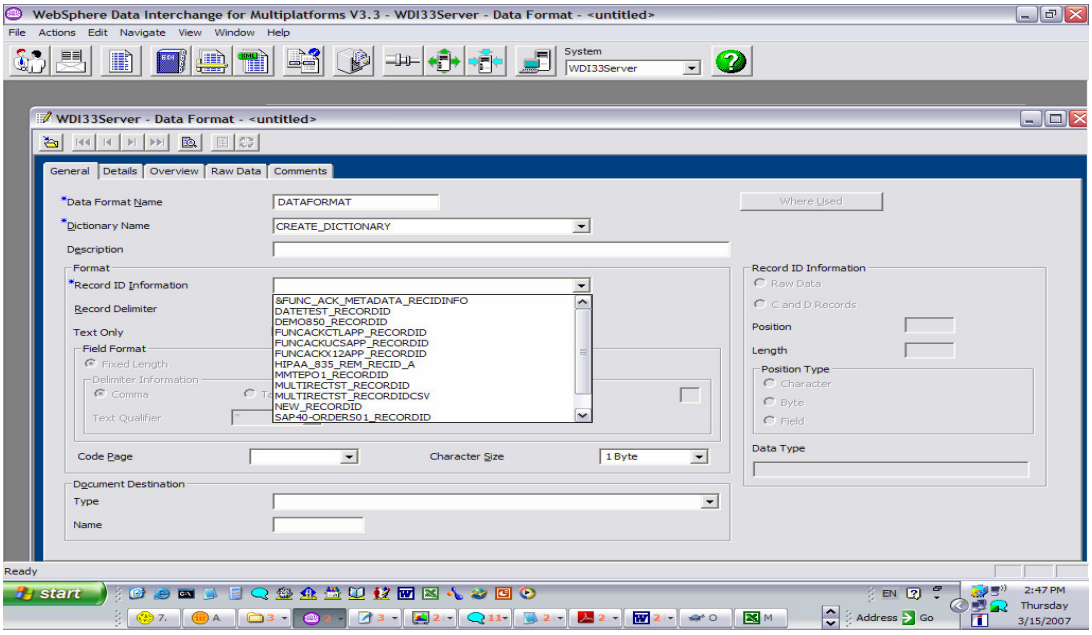
The Occurrence Information section indicates the maximum number of times a Loop repeats. The Unlimited Repeat flag indicates there is no maximum number and the repeats are not limited.

Creating Data Formats



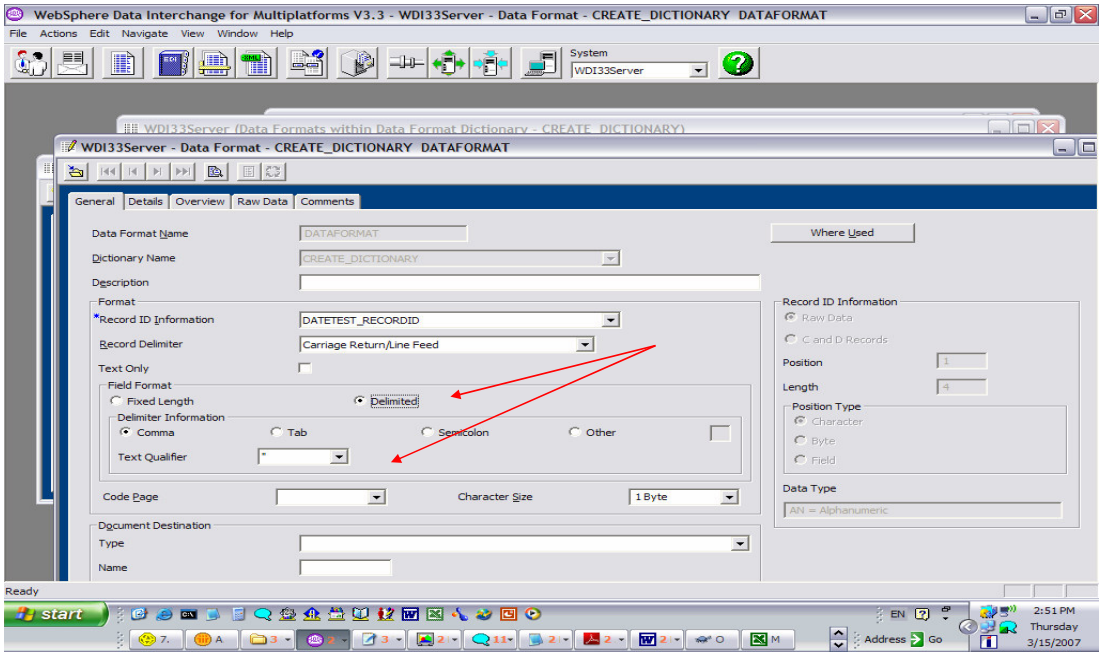
When all the fields, structures, records and loops have been defined, you can define the Data Format Definition.

Creating Data Formats



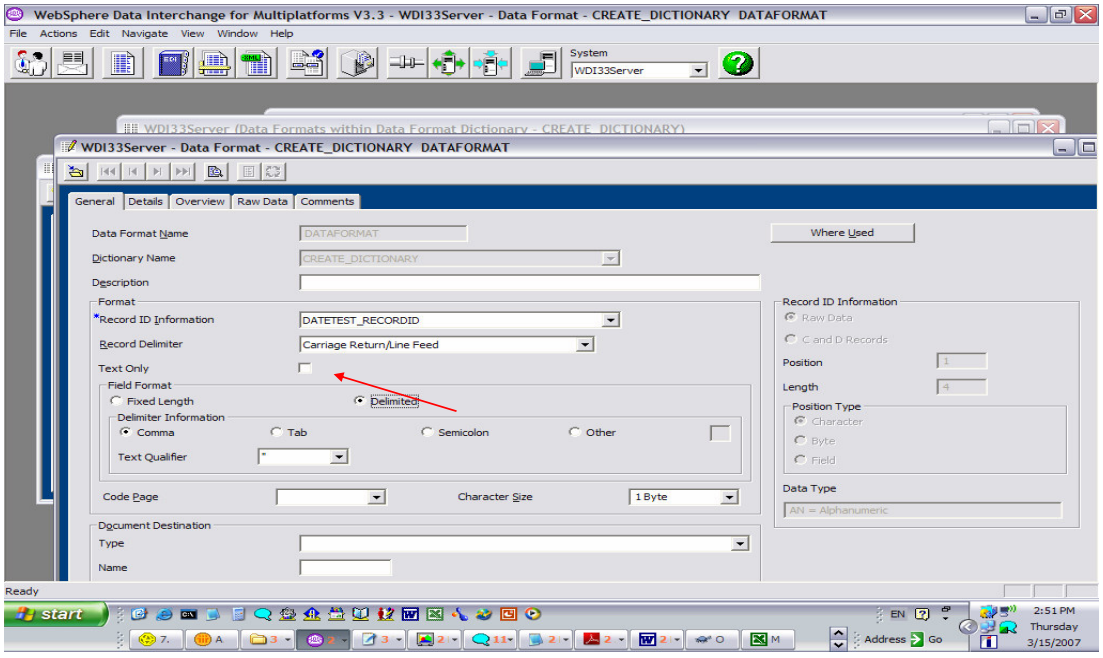
Enter the Data Format Name and Dictionary name and select the Record ID Information. The Record ID Information is a required field and if you have not defined this object you can either cancel here and create the object or select something in the list and change it later after you have created the object.

Creating Data Formats



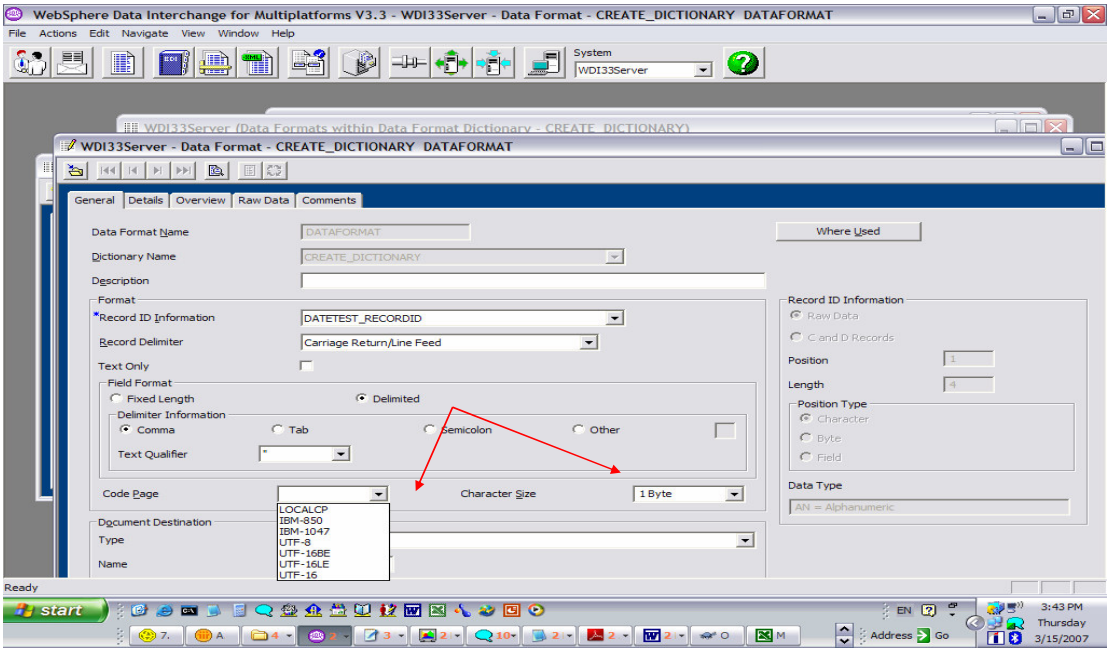
The Record Delimiter field indicates the input records are delimited with Carriage Return and Line Feed or New Line. A Record delimiter must be entered for input data that contains delimiters between values for example comma separated values. The delimiter and text qualifier are selectable. If the Records in a Data Format will contain delimited fields, then specify the delimiter to be used to separate fields and specify the text qualifier that will be used to enclose text fields. Records always contain only fixed length fields when the Data Format is used in Send or Receive Maps

Creating Data Formats



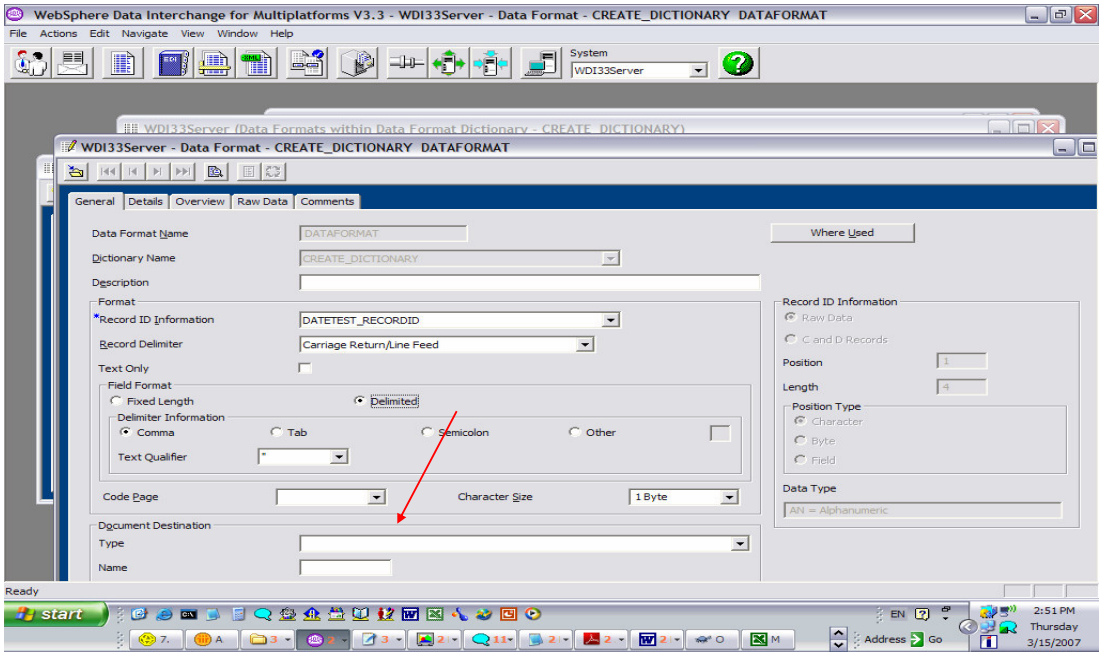
The Text Only flag indicates the input data does not contain any binary type values. During translation, there is additional processing required for binary type fields and this flag will skip this processing. This field is not used for Send and Receive processing.

Creating Data Formats



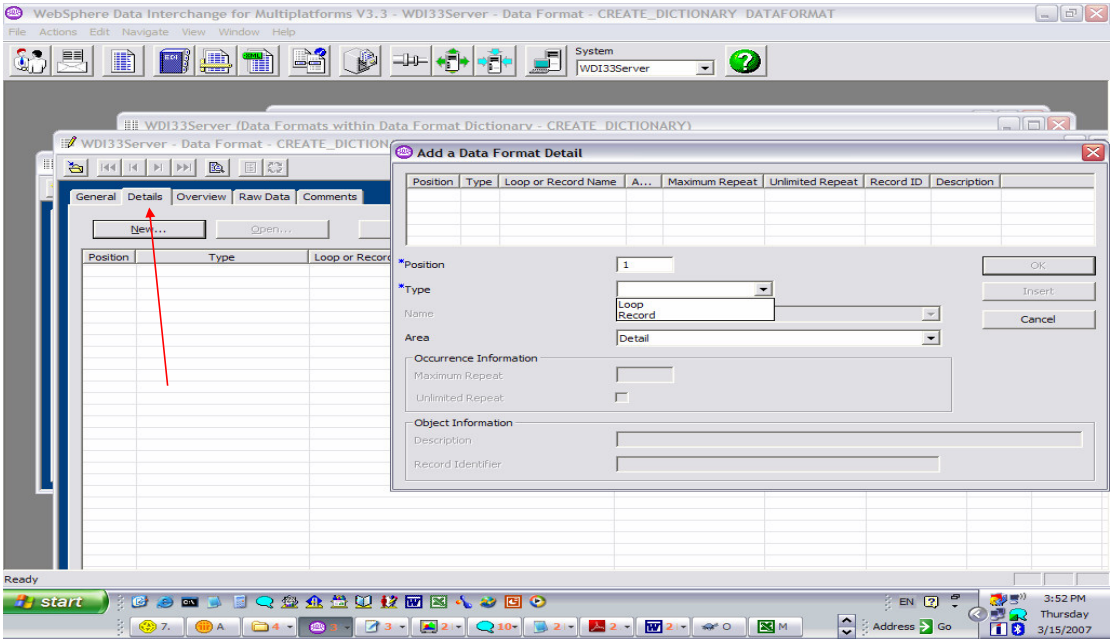
The code page for the data format can be selected from the drop down list along with the Character Size. For Example UTF-16 code page would have a Character size of 2 Bytes.

Creating Data Formats



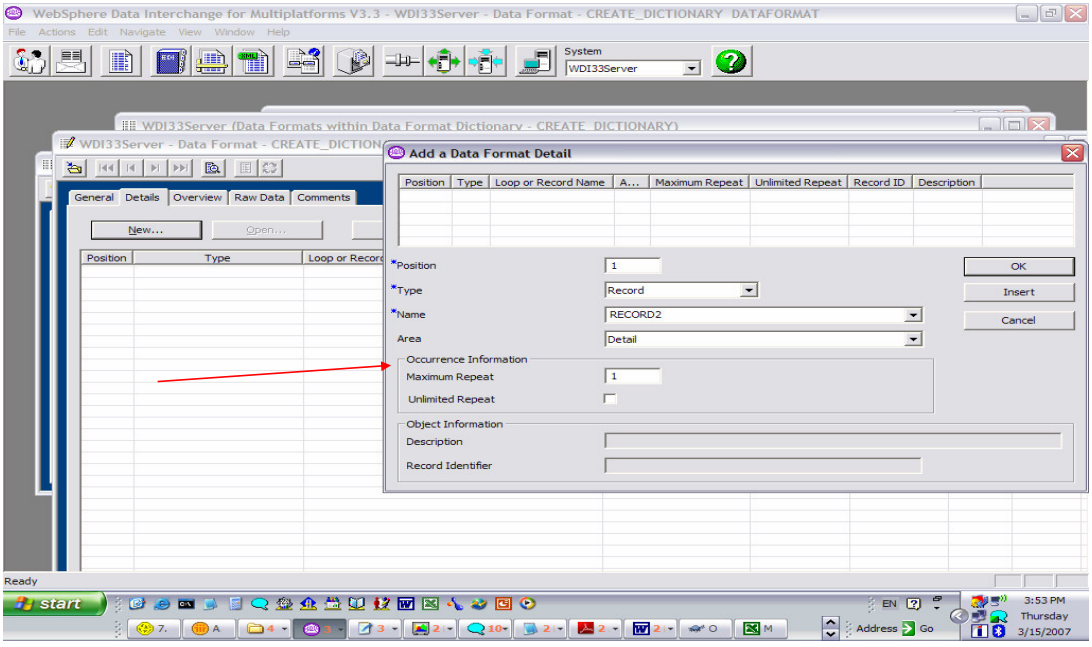
Specify the default output destination for documents created using the Data Format in the Document Destination area.

Creating Data Formats



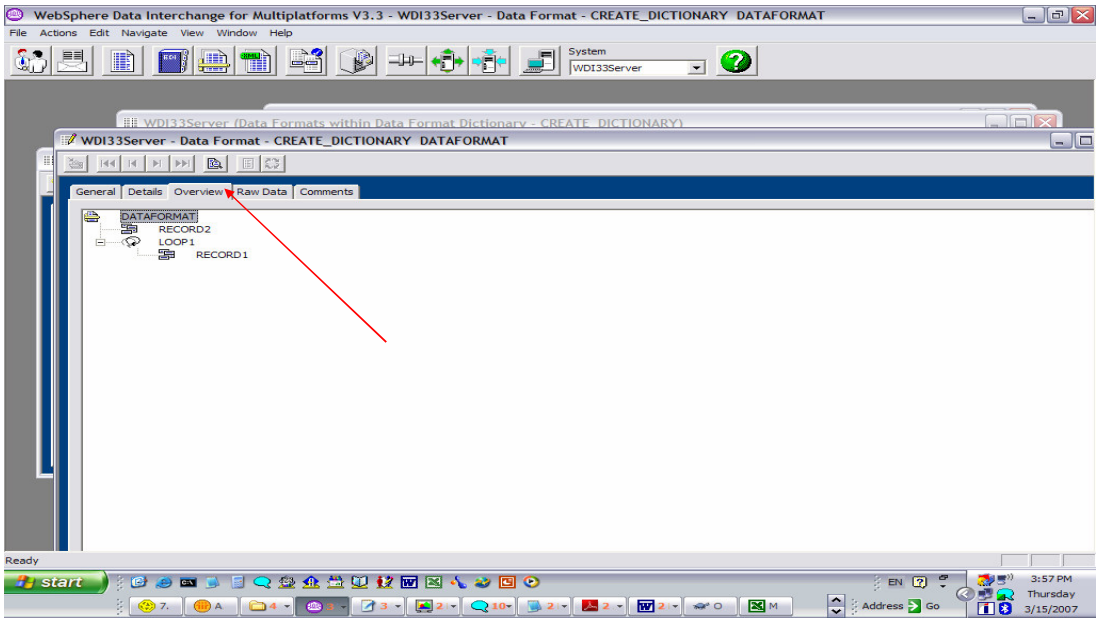
The Details Tab is where the Records and Loops that define your application are defined.

Creating Data Formats



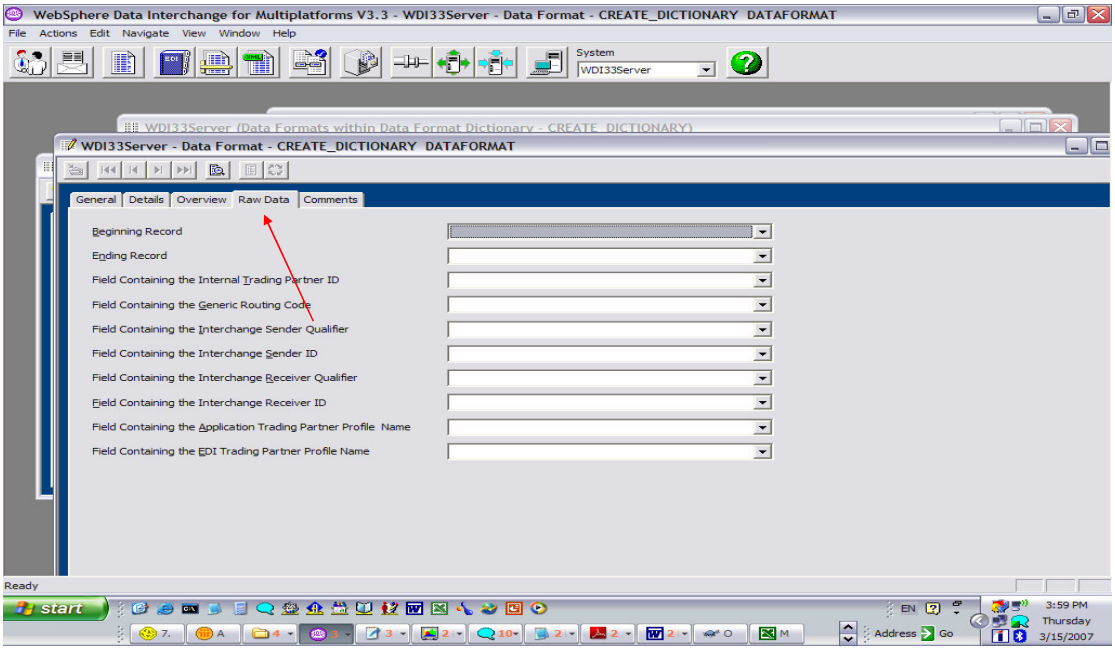
The Area selection can be either Header, Detail, or Trailer. The Occurrence Information area is where you define how many times the Record or Loop repeats and you may select Unlimited Repeat.

Creating Data Formats



The Overview Tab will display the Data Format in a Tree representation which is similar to what you will see during the mapping process.

Creating Data Formats



This Raw Data tab page on the Data Format Editor allows you to indicate which fields will contain important information that can be used in a translation by the WebSphere Data Interchange Server. The information is used only when the Data Format is Raw Data format. You do not need to enter anything on this tab page if the Data Format is C and D records format. Any fields specified on this tab page must occur only once within the Data Format.

Identify either the first or last record within the Data Format.

Select the name of the Data Format Field that will contain the internal trading partner ID if the Data Format contains such a Field. The internal trading partner ID is a value you use to identify your trading partner.

Optionally select the name of the Data Format Field that contains a generic routing code. The generic routing can be used to identify a generic Send Map Usage if a specific Send Map Usage cannot be located. This field is used only in Send Map translation.

If your document will contain the interchange sender qualifier and interchange sender ID, identify the Fields that will contain those values.

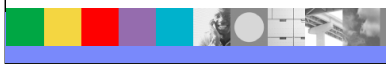
If your document will contain the interchange receiver qualifier and interchange receiver ID, identify the Fields that will contain those values.

If your document will contain the name of the Trading Partner profile that represents the application trading partner, identify the Field that will contain that value.

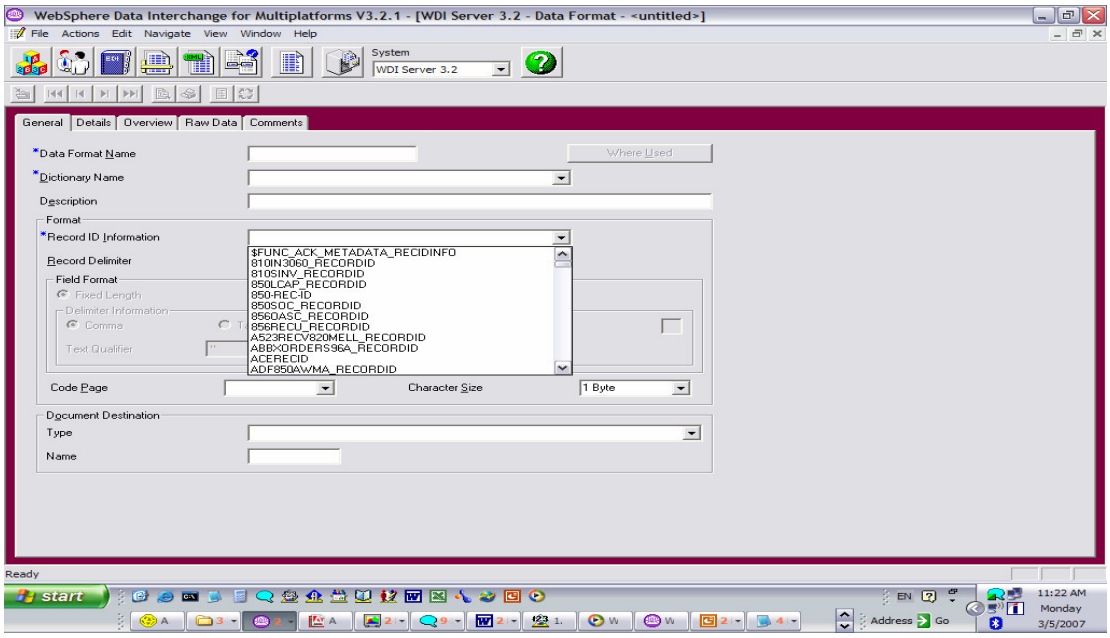
If your document will contain the name of the Trading Partner profile that represents the EDI trading partner, identify the Field that will contain that value.

Section

Record Id Information Object



Creating Data Formats



The Record Id information is selectable when you create a Data Format. This means that all records defined in the data format must conform to the same record Id position and length. With SAP Idocs for example, the header record is normally 10 bytes, but detail records may have record ids up to 30 characters.

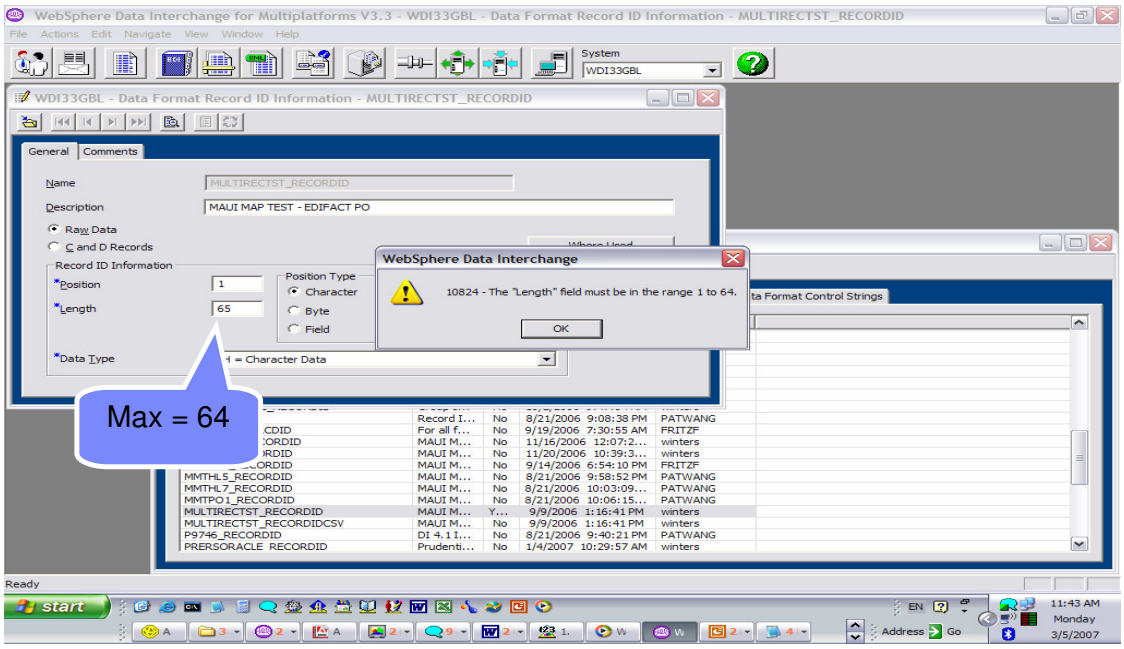
Creating Data Formats

- WDI V3.3
  - Allow varying/multiple record ids in a Data Format
  - Allow the records in the Data Format Definition to have different lengths for the record ids.
  - The maximum length extended to 64
  - WDI 3.2 control strings converted to WDI 3.3 control strings

With WDI Version 3.3 record identification is selectable for each record definition. This allows the value for Default Record Id Information and New Record Id Information. There is a Field Selection Option which allows an override for the Record Id Information. Users will be allowed to enter the Record Id value after selecting each Field that will be part of the Record Id Value. And the maximum length has been extended to 64 characters.

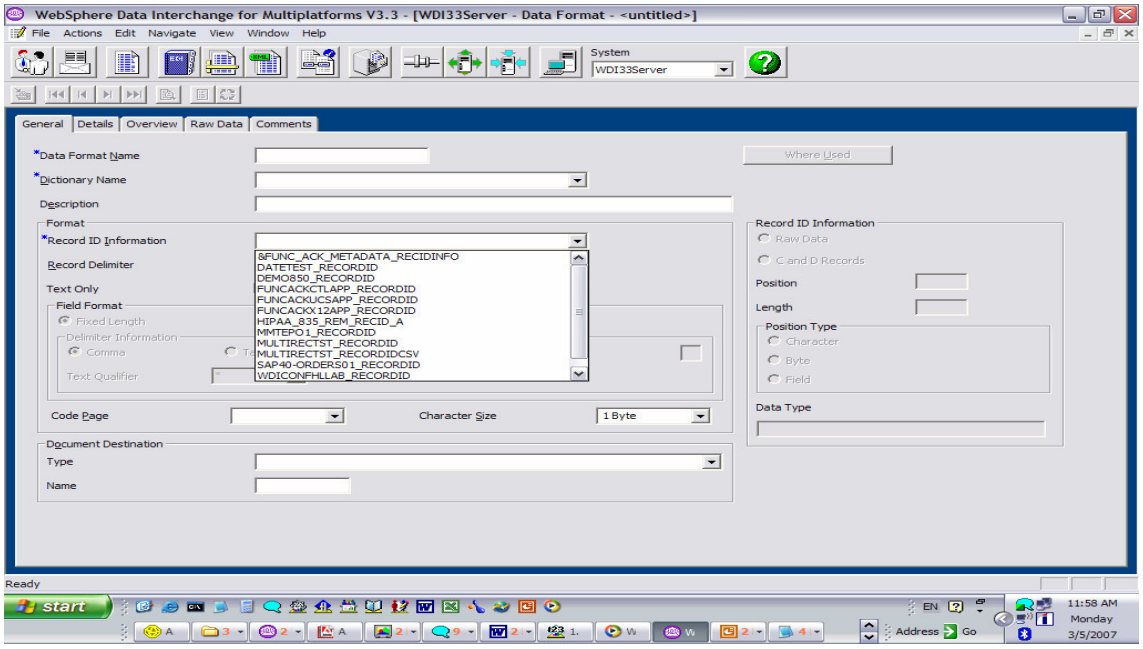
With this enhancement version 3.2 control strings will be converted to a version 3.3 layout for backward compatibility. The Control String is readable using an abstract layer which will load internal Control String to be used for processing.

Creating Data Formats



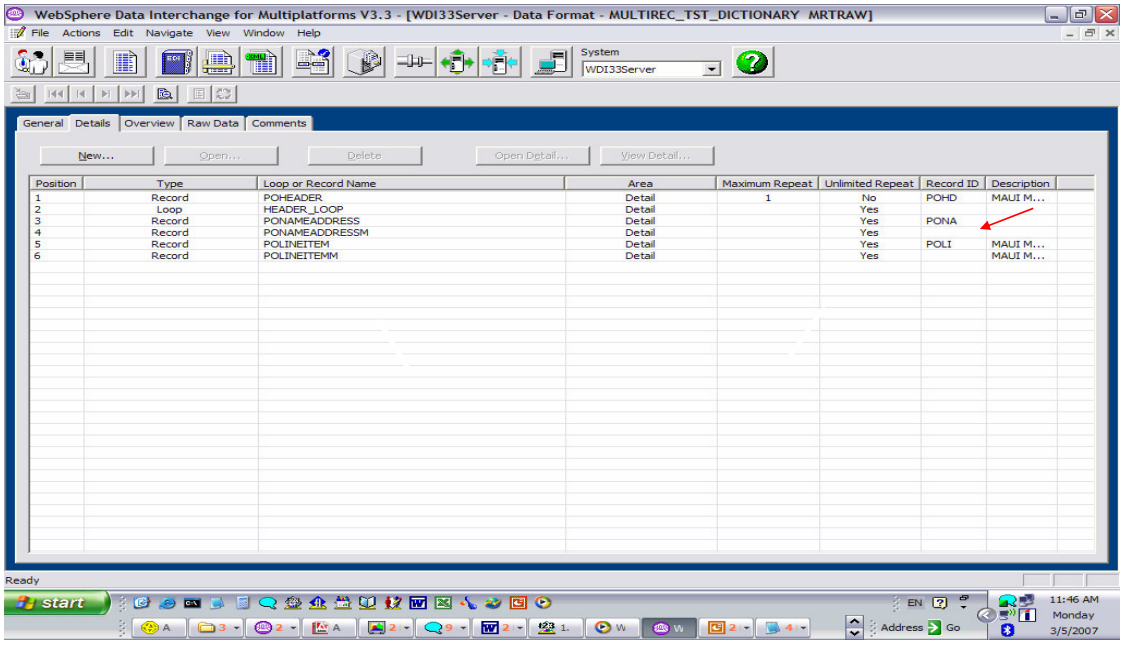
The maximum length is 64 characters for record Id length.

Creating Data Formats



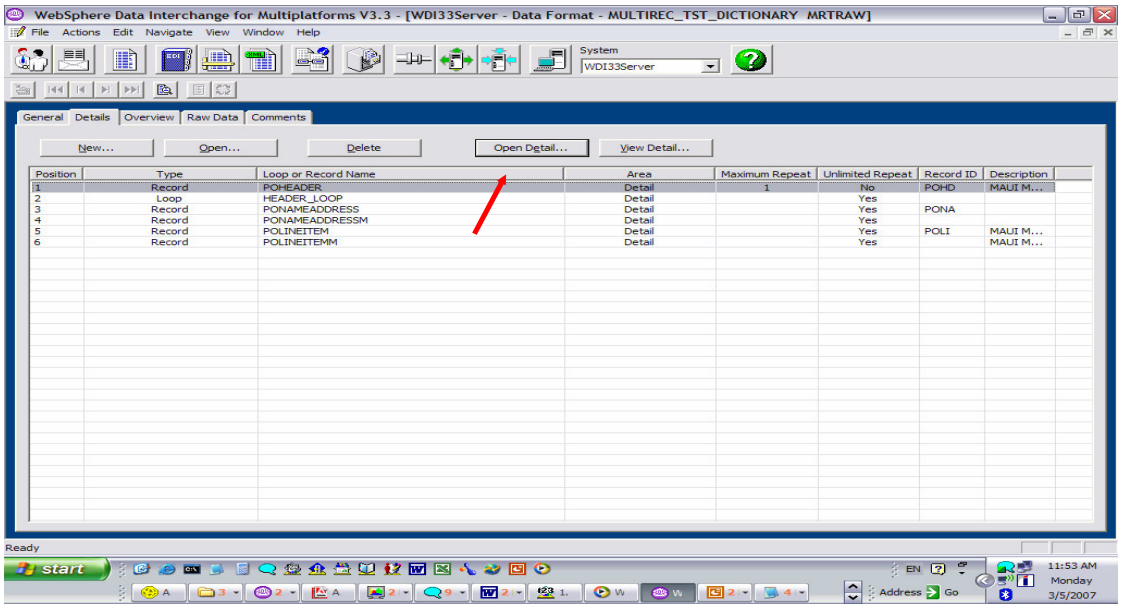
When you create a Data Format the Record Id Information is a mandatory field. The selection at the Data Format level will become the default Record Id information for all records defined in the Data Format.

Creating Data Formats



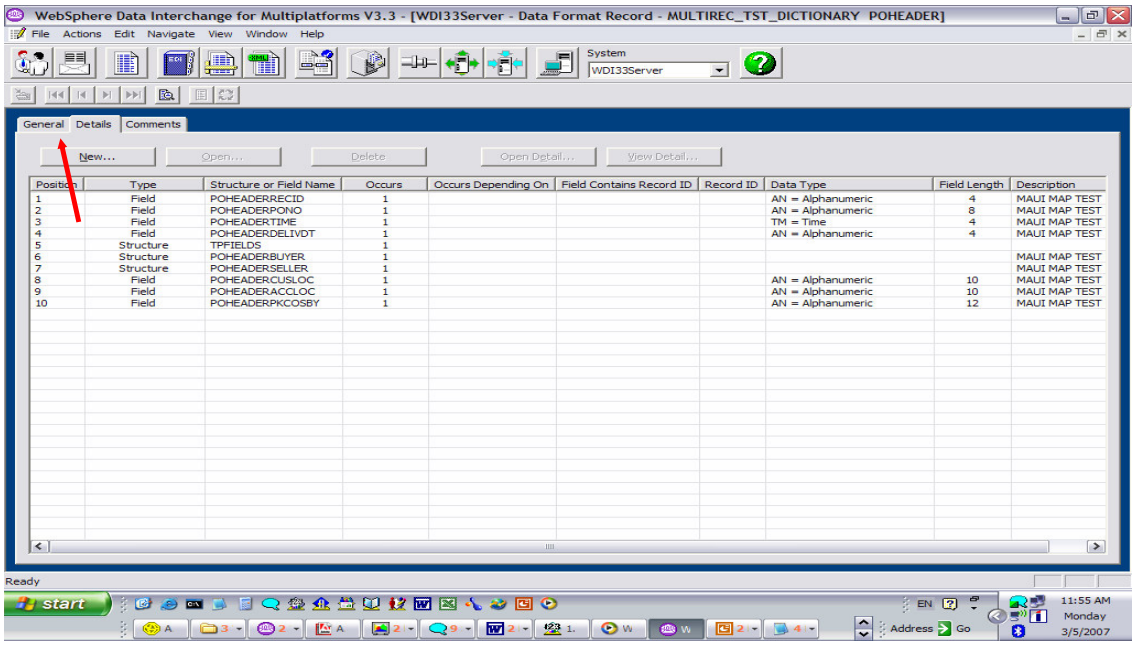
Note that some of these record definitions do not have a record Id value defined.

Creating Data Formats



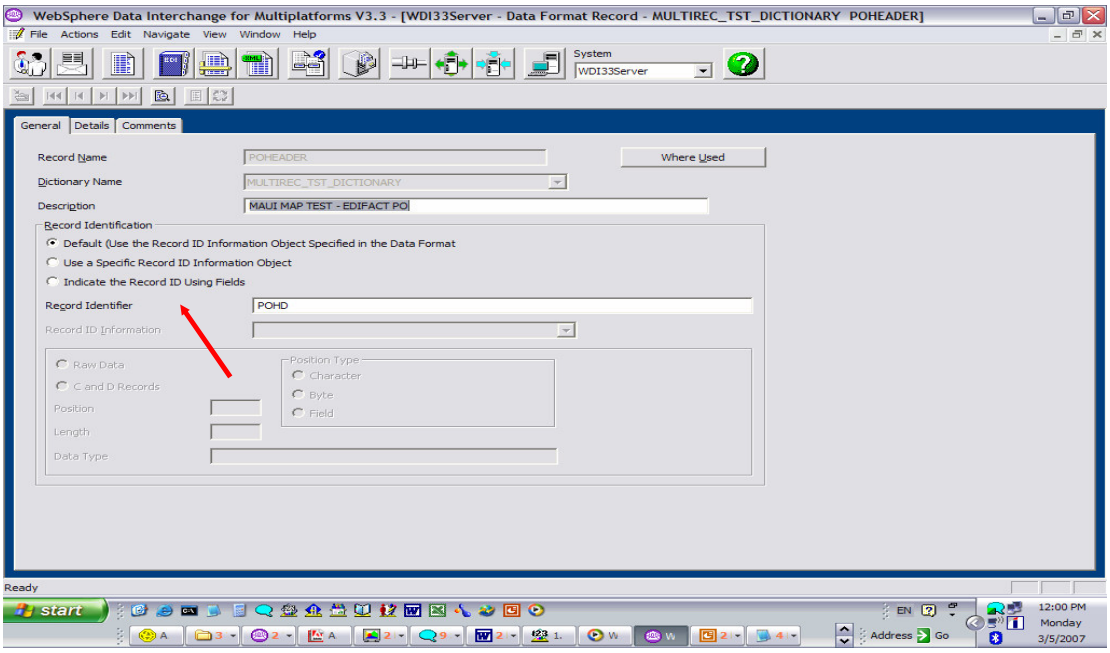
The Open Detail button will allow you to update all information including the Record Identifier.

Creating Data Formats



To change the Record Id Information, move to the General Tab.

Creating Data Formats



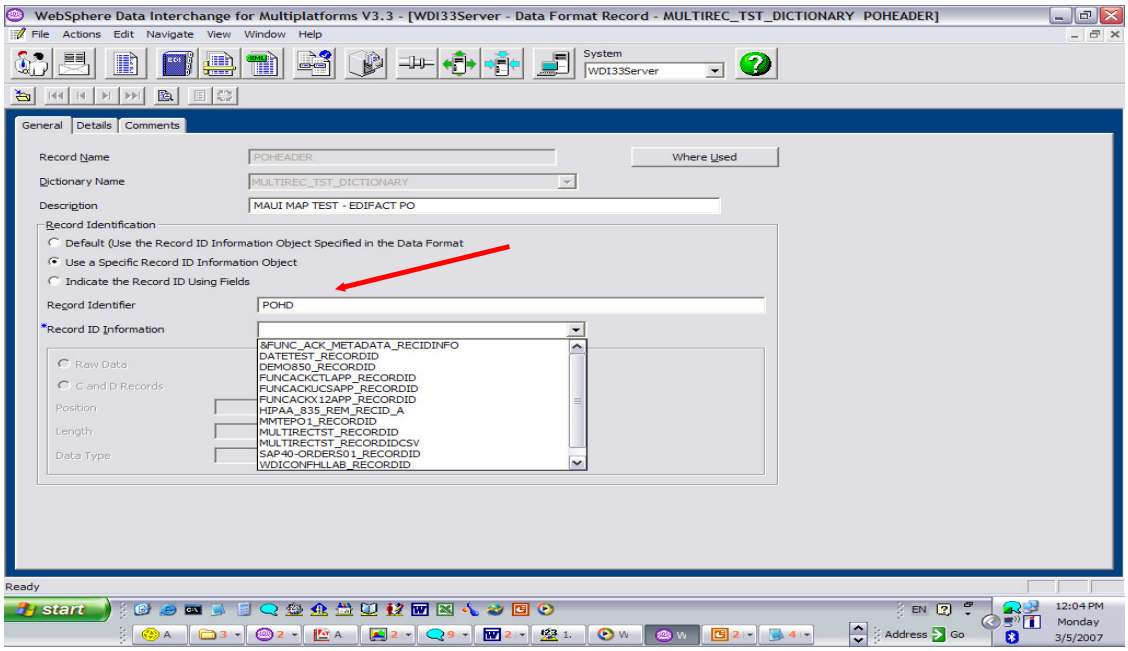
You can select to use the Default record Id Information defined for the data format, use a specific Record Id Information Object, or identify the Record Id using multiple fields in the record.

Creating Data Formats

- Default - This record uses the record Id from it's parent Data Format.
- Record ID Information - Select from a list of record ID's.
- Record ID using Fields - Identify a field or multiple fields as the record ID.
- The first radio button "Default" is highlighted by default when the user navigates to this Window.

If the default is used, you can enter the value of "Record Identifier text field" in a corresponding text field for this radio button and the size of this text field is limited to 64 characters. You can also select a different Record ID Information object and enter the "Record Identifier text field". Another option is to Select multiple fields that when combined will uniquely identify each record.

Creating Data Formats



When you select Use a Specific Record Id Information Object, the drop down list becomes active.

Creating Data Formats

Example:

Record ID Info: Record Id offset = 1 with length = 17
POHEADER: Record Id offset = 1 with length = 10
Default Record Id offset is 1 and length is 17.
Records in the Source data contain records

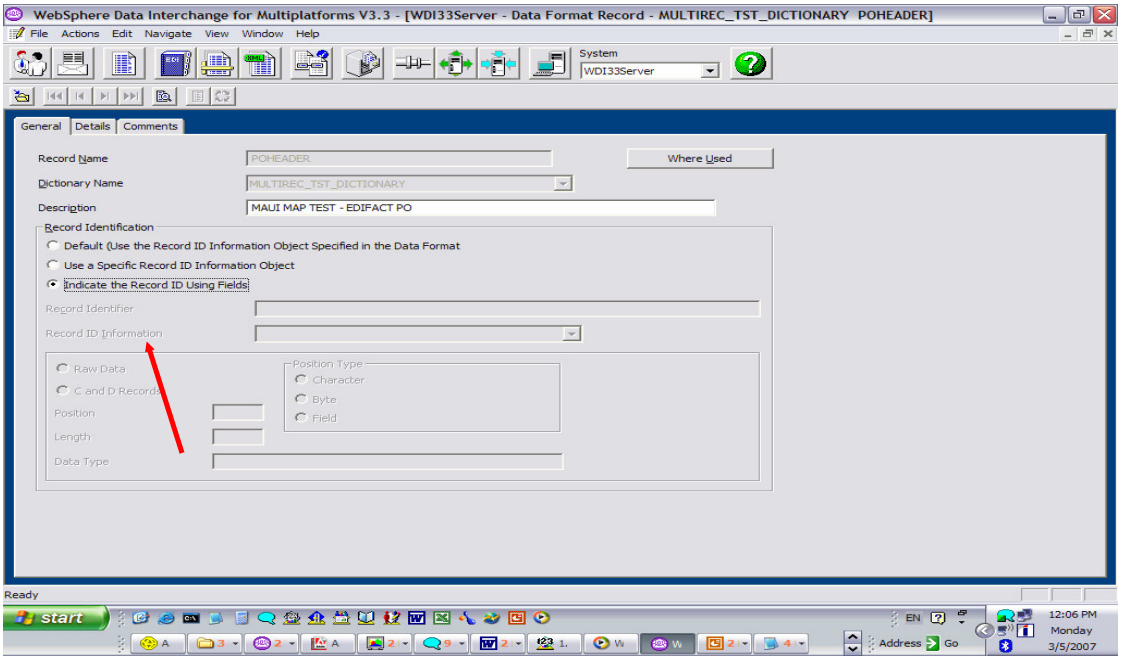
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0
POHEADER (data)
PONAMEANDADDRESS1 (data)
PONAMEANDADDRESS2 (data)

Records defined for the data format definition have the following values:

POHEADER
PONAMEANDADDRESS1
PONAMEANDADDRESS2

In this example, POHEADER can either use a Field or a different Record Id Information.

Creating Data Formats



When you select to define the Record Id using fields, the record Id value and the Record Id Information become protected.

Creating Data Formats

Example:

Record Id is offset 1 with length 13. There are 2 possible record layouts for PONAMEADDRESS. Another field in the PONAMEADDRESS record identifies the record layout.

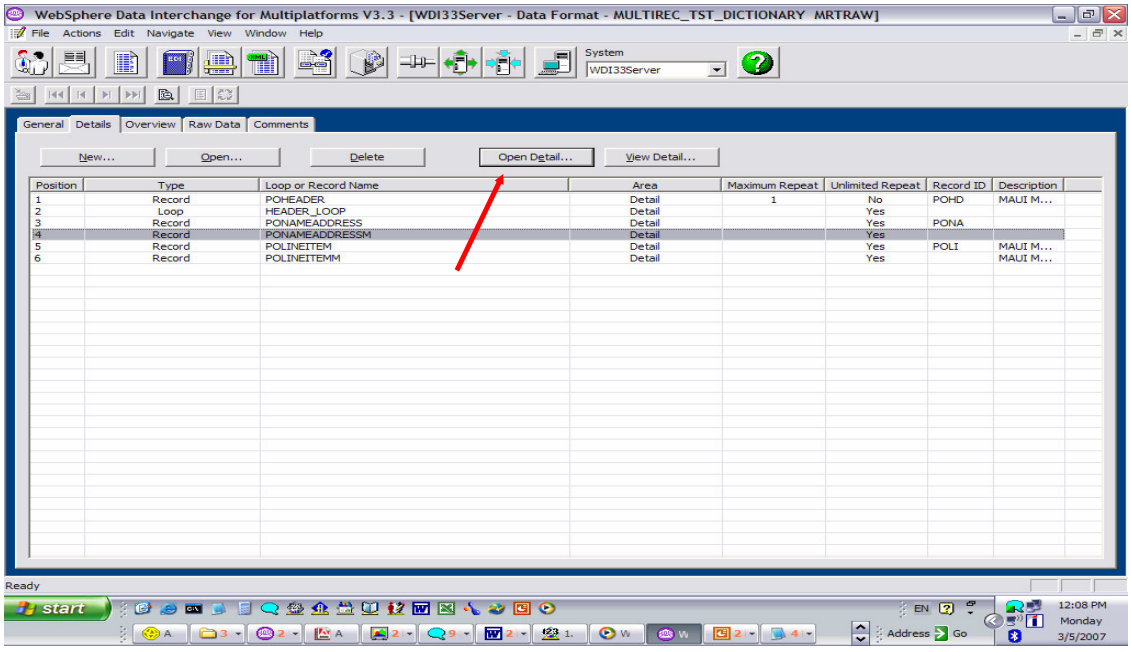
Record Id offset is 1 and length is 13.
Record Id offset for PONAMEADDRESS is 26 and length is 7.
Records in the Source data contain records

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2
POHEADER (data)
PONAMEADDRESS (data) LAYOUT1
PONAMEADDRESS (data) LAYOUT2

In this example, POHEADER Record defined for the data format definition can use the “Default Record Id Information” with the value “POHEADER”.

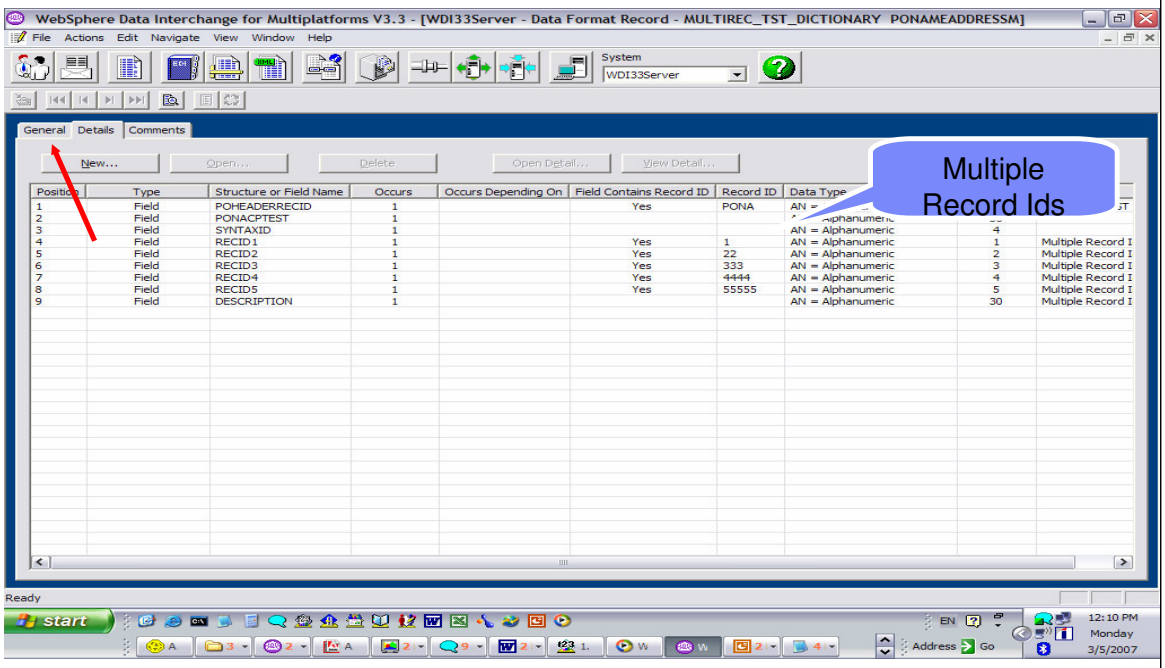
The PONAMEADDRESS record will become 2 records and will use 2 fields to define the record Id value. The field containing the value PONAMEADDRESS, for example Field1, and the field containing the values LAYOUT1 and LAYOUT2, for example Field 2, will be selected. Both records will contain a record Id value in Field 1. Field 2 will contain the values that make them different layouts or records.

Creating Data Formats



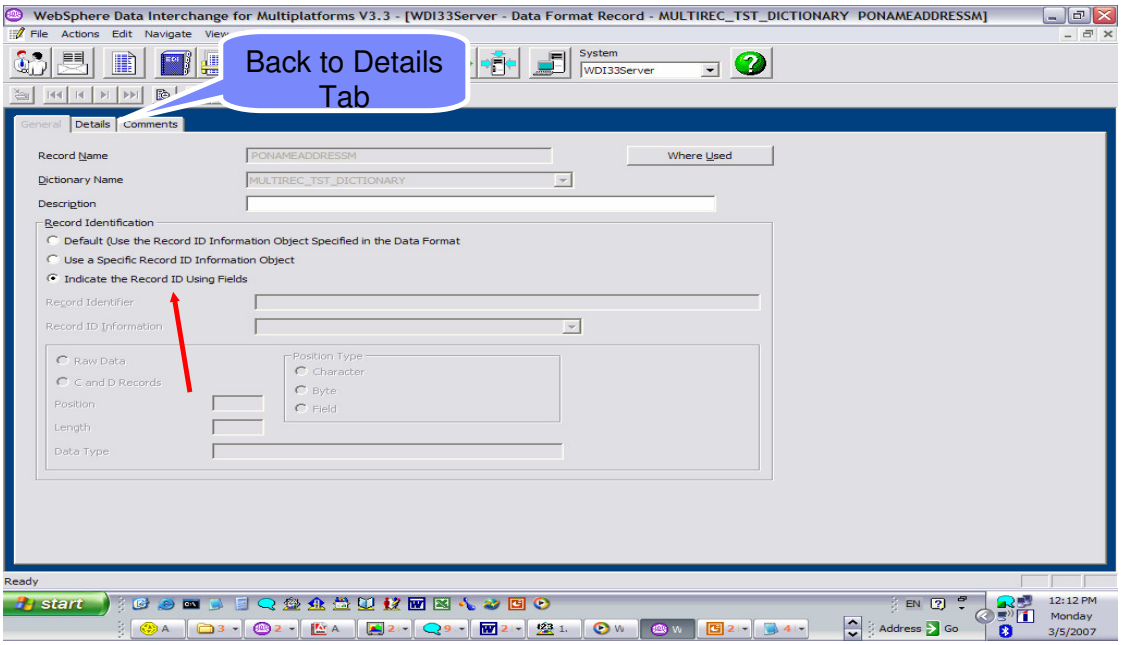
The PONAMEADDRESS record does not have a Record Id value, this may indicate record Ids defined as fields.

Creating Data Formats



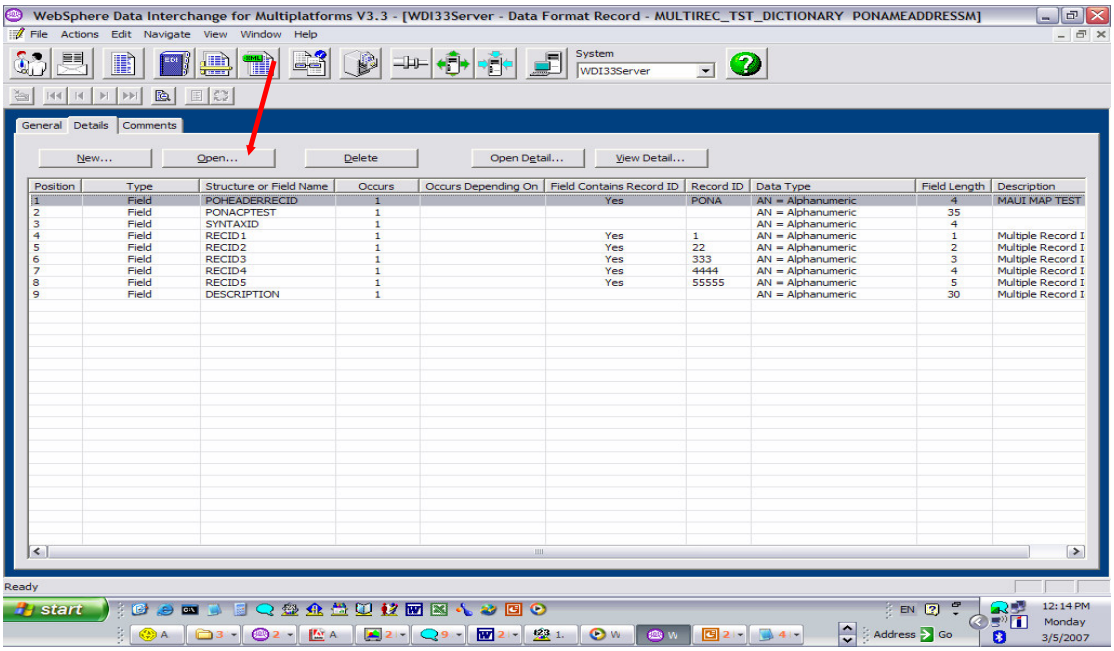
This record has multiple record ids defined. Lets move to the General Tab.

Creating Data Formats



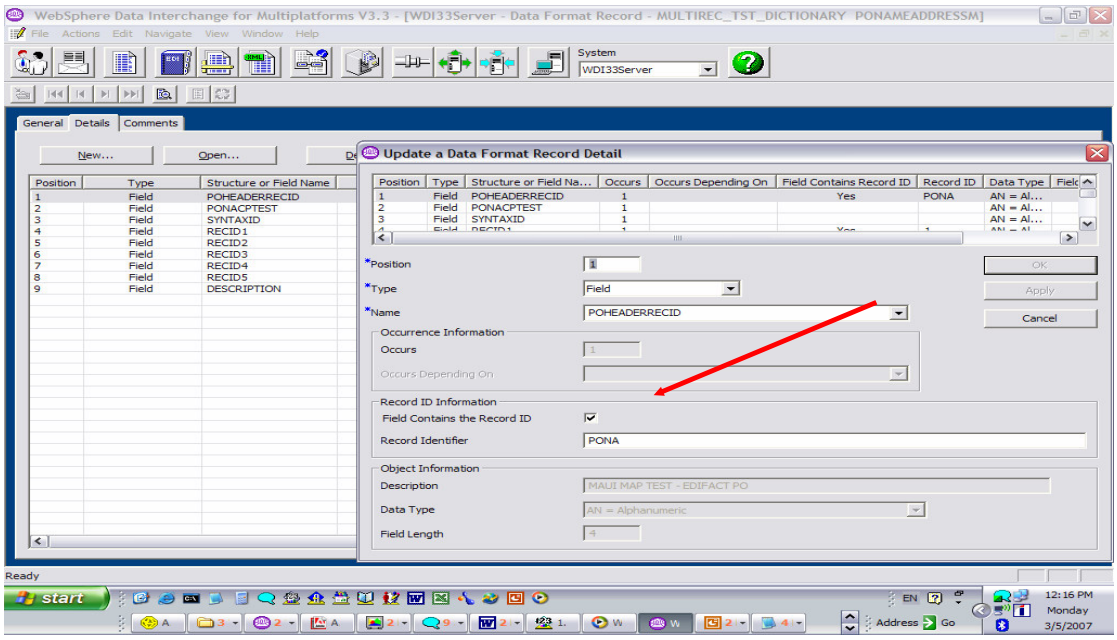
The Record Identification has Record Ids using fields selected.

Creating Data Formats



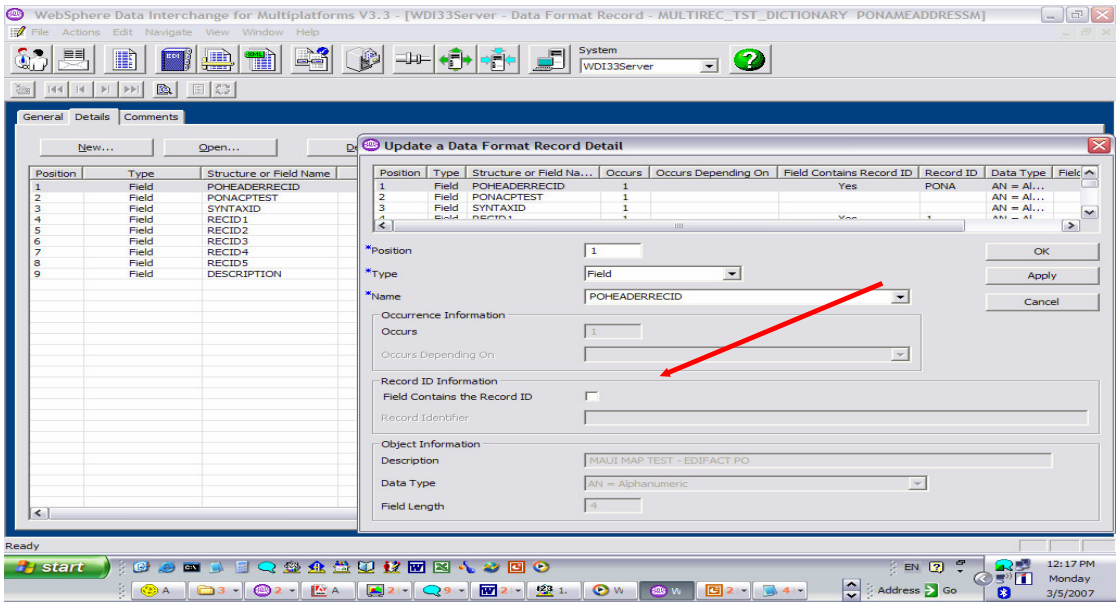
You can see which fields have been defined as part of the record Id in the Field Contains Record Id column. If we open the POHEADERRECID field, we can update the information for this Record Id.

Creating Data Formats



To define a field as part of the record id select the flag “Field Contains the Record ID” and enter the Record Id value.

Creating Data Formats



To remove a field as part of the record id un-select the flag “Field Contains the Record ID”.

Creating Data Formats

- Restrictions
  - Record Ids may not be defined as fields within a structure.
  - Record ID Data Type restrictions: Record Id values are restricted to character-based fields.
  - Using Occurs Depending Structures, Record Id Selection must occur before the variable length structure within the record. NOTE: This is also a current restriction for WDI 3.2.

Restrictions include Record Ids may not be defined as fields within structures and Record Id data type must be a character-based data type. Using Occurs Depending structures, the Record Id must occur before the occurs depending structure.

Reference

- More information can be found in the WDI V3.3 Mapping Guide

More information can be found in the WebSphere Data Interchange Version 3.3 Mapping Guide.

Template Revision: 04/25/2006 11:09 AM

Trademarks, copyrights, and disclaimers

The following terms are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both:

IBM	CICS	IMS	WMQ	Tivoli
IBM(logo)	Cloudscape	Informix	OS/390	WebSphere
e(logo)business	DB2	iSeries	OS/400	xSeries
AIX	DB2 Universal Database	Lotus	pSeries	zSeries

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are registered trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a registered trademark of Linus Torvalds.

Other company, product and service names may be trademarks or service marks of others.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

Information is provided "AS IS" without warranty of any kind. THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, express or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. All customer examples described are presented as illustrations of how those customers have used IBM products and the results they may have achieved. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Note to U.S. Government Users - Documentation related to restricted rights-Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract and IBM Corp.