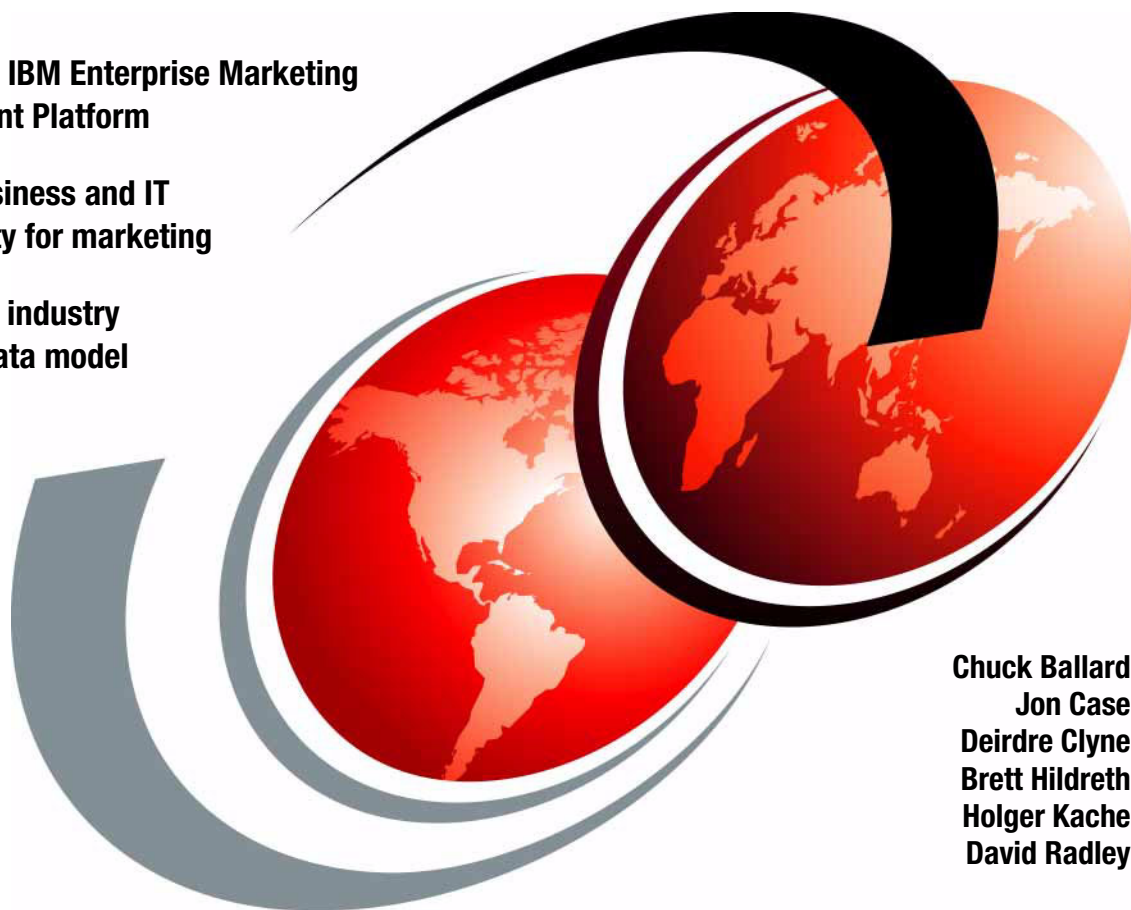**IBM**

Enhance Inbound and Outbound Marketing with a Trusted Single View of the Customer

Build on an IBM Enterprise Marketing Management Platform

Include business and IT functionality for marketing

Use an IBM industry standard data model

Chuck Ballard
Jon Case
Deirdre Clyne
Brett Hildreth
Holger Kache
David Radley

Redbooks

IBM

International Technical Support Organization

**Enhance Inbound and Outbound Marketing with a
Trusted Single View of the Customer**

March 2014

Note: Before using this information and the product it supports, read the information in "Notices" on page vii.

First Edition (March 2014)

This edition applies to IBM® InfoSphere® Master Data Management (MDM) Enterprise Edition, version 11.

Contents

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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 and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

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

DataStage®	InfoSphere®	Smarter Commerce™
DB2®	MetaBroker®	SPSS®
developerWorks®	PureData™	Unica®
IBM®	Rational®	Velocity™
IBM PureData™	Redbooks®	WebSphere®
IBM Watson™	Redbooks (logo) ®	

The following terms are trademarks of other companies:

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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

Preface

The term *customer centricity* is used with more frequency in executive board rooms and strategic planning meetings, within organizations that span all industries and all sizes. The essence of customer centricity is this: To thrive, organizations must understand their customers at a personal, individual level, and organizations must use that understanding to transform the way they run their business.

The importance of customer centricity might be greatest within a marketing organization. Traditional mass-market approaches to marketing are less effective today. Reaching consumers *en masse* through just one medium is more difficult because of increases in the number and variety of marketing channels. Consumers expect and reward more personalized and consistent experiences across channels and points of interaction. There is ongoing commoditization of products, promotions, and offers; competing on price alone is not sustainable because of low margins, so organizations must compete on experience and service too. To adapt to these changes and stay ahead of competition, organizations must evolve to use a customer-centric approach to marketing, which includes breaking down information silos, segmenting and analyzing customers at an individual level, and engaging consumers in interactive marketing.

This IBM® Redbooks® publication illustrates how businesses can use IBM software to achieve customer centricity in their marketing organization. The book provides a blueprint for breaking down information silos to establish a foundational cross-channel 360-degree view of individual consumers. It describes how you can use that 360-degree view to improve both your outbound and interactive marketing across all channels.

This book is applicable to both a marketing business audience who wants to understand how IBM can help transform their marketing management across the enterprise and an IT audience who wants to understand how IBM can help provide the data and information foundation to support that marketing transformation.

Authors

This book was produced by a team of specialists from around the world working with the International Technical Support Organization in San Jose, CA.

Chuck Ballard is a Project Manager at the International Technical Support organization, in San Jose, California. He has over 35 years of experience, holding positions in the areas of product engineering, sales, marketing, technical support, and management. His expertise is in the areas of database, data management, data warehousing, business intelligence, and process re-engineering. He has written extensively about these subjects, taught classes, and presented at conferences and seminars worldwide. Chuck has both a Bachelor's degree and a Master's degree in Industrial Engineering from Purdue University.

Jon Case is a Product Manager for the IBM InfoSphere® Master Data Management software product, based in Raleigh, North Carolina. He has over 15 years of experience in IT, including 9 years experience with master data management in roles spanning implementation services, sales engineering, and product management. In his role, Jon works with IBM clients on strategies for bringing the value of master data management from the infrastructure or back-office level to the front-office level, putting master data into the hands of business users and applications to enable more effective business solutions. Jon has a Bachelor's degree in Engineering from the University of Minnesota.

Deirdre Clyne is an Information Integration Technical Sales Specialist in North America. She has over 20 years of experience in Information Management in a variety of roles including sales, support, technical sales, and development. Her areas of expertise include data modeling, metadata management and data governance. She holds a degree in Electronic Engineering from Trinity College Dublin.
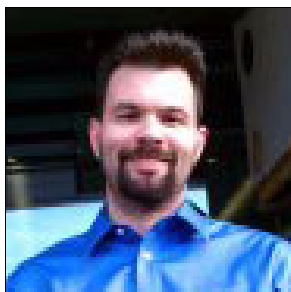
Brett Hildreth has spent nearly 15 years focusing on the Enterprise Marketing Services and Technology industry. During that time he has worked with competitive marketing technologies, marketing services organizations, and clients around the globe across nearly every industry vertical market. Since joining IBM, by way of the Unica® acquisition, Brett has supported Enterprise Marketing Management (EMM) sales efforts in the capacity of a World Wide Client Solutions Professional, and more recently, as a Senior Strategist within the EMM organization, focused on EMM Innovation and Integration efforts.

Holger Kache is an IBM InfoSphere Solution Architect at IBM Germany. He has 14 years of experience in the industry and holds a Master's in Computer Science from the Technical University of Cottbus, Germany. Holger spent 13 years of his career at the IBM Silicon Valley Labs to develop Information Management software. His expertise is in Data Integration and Governance and includes several research and industry publications and patents.

David Radley is an IBM InfoSphere Solution Architect in the IBM Hursley lab in the UK. He has 25 years of experience in IT, with the last 8 years in Information Management. In his role, David promotes and develops trusted predictive analytics and EMM solutions by putting master data at the heart. In 2013, he wrote an IBM developerWorks® article *Enhancing IBM Enterprise Market Management solutions with master data. A trusted base for EMM*, which is at the following location:

http://ibm.co/1dafICE

David holds a degree in Physics from Birmingham University.

Other Contributors:

We thank others who contributed to this IBM Redbooks publication, in the form of written content, subject expertise, and support.

From IBM locations worldwide

► John Baldwin: Product Consultant MDM Server, IBM Software Group, Information Management, London, UK
► Mandy Chessel: FEng, CEng, FBCS, Distinguished Engineer, Master Inventor, Chief Architect for InfoSphere Solutions, IBM Software Group, Information Management, Hursley Park, Winchester, UK
► Mike Cobbett: Master Data Management Tools Architect, IBM Software Group, Information Management, Hursley, GB
► Bhavani Eshwar: Software Architect, Master Data Management, IBM Software Group, Information Management, Bangalore KA, IN
► Robert A. Alfieri: Senior Enterprise Architect, Worldwide Sales, IBM Software Group, San Antonio, Texas, USA
► Elaine Hanley: Information Architect, Data Warehousing, IBM Software Group, Information Management, Dublin, Ireland
► Jay Henderson: Global Strategy Program Director, Enterprise Marketing Management, IBM Software Group, Waltham MA, USA
► Vinnie Mastro: Development Manager, Information Server Metadata, IBM Software Group, Information Management, Littleton MA, USA
► Joe O'Sullivan: Solution Sales Manager, IBM Industry Models, IBM Sales and Distribution, Software Sales, Dublin, Ireland
► Patrick O'Sullivan: Senior Technical Staff Member, IBM Industry Models, IBM SWG Information Management, Dublin, Ireland
► Lena Woolf: Senior Technical Staff Member, InfoSphere MDM, IBM Software Group, Information Management, Markham ON, CA

From the International Technical Support Organization

► Mary Comianos, Publications Management
► Ann Lund, Residency Administration
► IBM Redbooks Graphics Support Team
► Diane Sherman, Editor

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

► Use the online **Contact us** review Redbooks form found at:

   ibm.com/redbooks

► Send your comments in an email to:

   redbooks@us.ibm.com

► Mail your comments to:

   IBM Corporation, International Technical Support Organization
   Dept. HYTD Mail Station P099
   2455 South Road
   Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction

The term *customer centricity* is used with more frequency in executive board rooms and strategic planning meetings, within organizations that span all industries and all sizes. The exact definition used from one organization to the next might vary slightly, but the essence is this: To thrive, organizations must understand their customers at a personal, individual level, and organizations must use that understanding to transform the way they run their business.

The importance of customer centricity might be greatest within a marketing organization. Traditional mass-market approaches to marketing are less effective today for many reasons:

► Increases in the available media and market channels means that it is harder to reach consumers with just one campaign through one medium.

► Consumers have more choices, alternatives, and information, which means they have more power. They expect and reward more personalized and consistent experiences across channels and points of interaction.

► Commoditization of products, promotions, and offers is ongoing. Competing on price alone is not sustainable because of low margins, so organizations must compete on experience and service also.

To adapt to these changes and stay ahead of competition, organizations must evolve to use a customer-centric approach to marketing:

– Break down organization silos. Traditional organization around product lines or sales channels is not efficient and will not yield the personalized, consistent experience consumers expect.

– Break down information silos. An enterprise-wide, cross-channel 360-degree understanding of your customers must be the foundation of your marketing execution.

– Become more granular in the way you analyze and segment consumers and the way you select and execute marketing offers to individuals.

– Engage consumers in interactive marketing. Take advantage of insights you can derive from both the historical information you know about a consumer and also the real-time context of each new interaction.

This book illustrates how organizations can use IBM software to achieve customer centricity in their marketing organization.

► Chapters 2 - 5 provide a blueprint for breaking down information silos and how IBM Information Management capabilities are used to establish the foundational cross-channel 360-degree view of individual consumers. These capabilities include master data management, data integration, metadata management, industry models, and workload-optimized analytics appliances.

► Chapters 6 - 7 describe how IBM Enterprise Marketing Management (EMM) capabilities are used to select and deliver the most relevant and personalized marketing offers to consumers. The chapters describe how to take the 360-degree view and use it to improve your marketing across all channels: analyzing data to find actionable insights, deciding on the best offer or communication, delivering engaging messages, measuring results, and capturing reactions and data to augment each customer profile. These chapters cover both outbound marketing and inbound interactive marketing.

► Chapter 8 highlights several next steps you might want to consider after you establish the foundation of trusted consumer information and EMM.

1.1 Establishing a trusted 360-degree customer view

Chapters 2-5 of this book describe how the IBM Information Management platform is used to integrate customer information across channels to establish a trusted 360-degree customer view. These chapters provide a detailed description that builds upon an IBM white paper (*Enhancing IBM Enterprise Market Management solutions with master data. A trusted base for EMM*), which is available at IBM developerWorks site:

http://www.ibm.com/developerworks/data/library/techarticle/dm-1303enhance/

Many organizations have some existing data warehouse or data store that is the closest approximation they can get to a 360-degree view of their customers. However, these existing data stores fall short in several ways:

► Incomplete: They do not always include customer and prospect information from every brand, line of business, or sales channel.

► Inaccurate: They use, at best, simple rules and deterministic logic to identify duplicate customer records, resulting in only a small fraction of the duplicates detected and consolidated into a single view.

► Inconsistent: They struggle to bring together information from disparate sources in a consistent model that is optimized for their industry or for marketing purposes.

► Workload-optimized: They cannot support both real-time demands of interactive marketing and also analytical demands of segmentation, analysis, and out-bound campaign marketing.

The subsequent chapters describe how the following IBM Information Management capabilities address each of those challenges:

► Master data management (MDM)

MDM uses a sophisticated probabilistic matching algorithm to accurately identify duplicate customer and prospect information from disparate data sources. MDM also uses business rules to bring together multiple duplicate records in a way that presents them as a trusted single unique view of the customer. MDM is optimized for real-time operations to add, update, search, and retrieve this core customer profile, which is critical for real-time interactive marketing. MDM also shares the core customer profile with a marketing repository where it is combined with richer transaction history for use in analytical decision-making.

- ► Industry models

  IBM industry models provide an industry-specific framework through which you can organize your customer information, both the core customer profile and the deeper transaction and activity details about your customer.

- ► Workload-optimized marketing repository

  The relevant customer information is housed in a marketing data repository appliance that is optimized for fast, easy analytical use, whether that is running segmentation analysis or pulling larger lists for outbound campaigns.

- ► Information integration

  Information integration capabilities can help to more easily move information through an information supply chain from the original sources of customer information, through MDM, and ultimately to the marketing data repository. IBM information integration capabilities include greatly simplified synchronization between MDM and the marketing data repository to minimize the time and effort required to establish the trusted 360-degree customer view. IBM also provides data governance capabilities as part of the information integration "information supply chain," including the ability to share a common glossary of business terms for all users, the ability to share metadata about your data architecture, data sources, and data consumers, and the ability to enforce data quality along the way.

1.2 Acting on the trusted 360-degree customer view

Chapters 6-7 of this book describe how IBM Enterprise Marketing Management (EMM) platform integrates marketing across channels to act on the trusted 360-degree customer view to power your marketing operations. The EMM platform allows marketers to integrate all their marketing efforts across all media types into one cohesive, coordinated marketing program.

The IBM EMM platform has a broad set of capabilities. EMM helps marketers do the following tasks:

- ► Collect data that enriches and augments what they already know about their customers and prospects: data such as what each customer has done on the company's digital properties or the history of marketing interactions and responses with each customer.

- ► Analyze their customer and prospect data to find new, actionable insights into their customer base that can increase the effectiveness of their marketing efforts.

- ► Decide on the next marketing action to take with each individual customer and prospect.

► Deliver the appropriate marketing message and capture the response. Those responses in turn can influence future messages, ensuring the customer dialogue remains interactive.

► Manage what is going on in the marketing organization, streamlining operations and improving marketing decisions by measuring results, tracking performance, and guiding future marketing investment decisions.

This book focuses on two specific EMM components that directly use the trusted 360-degree customer view to optimize decisions about marketing offers and automate the execution of those decisions.

► IBM Campaign: With IBM Campaign, marketers can plan, design, execute, measure, and analyze multi-wave, cross-channel, and highly personalized marketing campaigns. The recognized best-of-breed functionality and scalability of IBM Campaign plays a key role in defining and executing interactive dialogs with customers and prospects.

► IBM Interact: With IBM Interact, marketers can personalize, in real time, the experience of customers interacting with websites, call centers, and other inbound marketing channels. IBM Interact uses powerful, behavioral targeting analytics and marketer-defined business logic to deliver the optimal marketing message in each case.

1.3 Combining trusted information and Enterprise Marketing Management

The capabilities and components described fit together into an overall solution to create and act on the trusted 360-degree customer view. Subsequent chapters describe the use and configuration of each component. Before reading about those details, review the overall solution in Figure 1-1.

This book follows the supply chain of information that flows from line-of-business applications, through MDM and the data warehouse, to the marketing repository, as depicted in Figure 1-1.
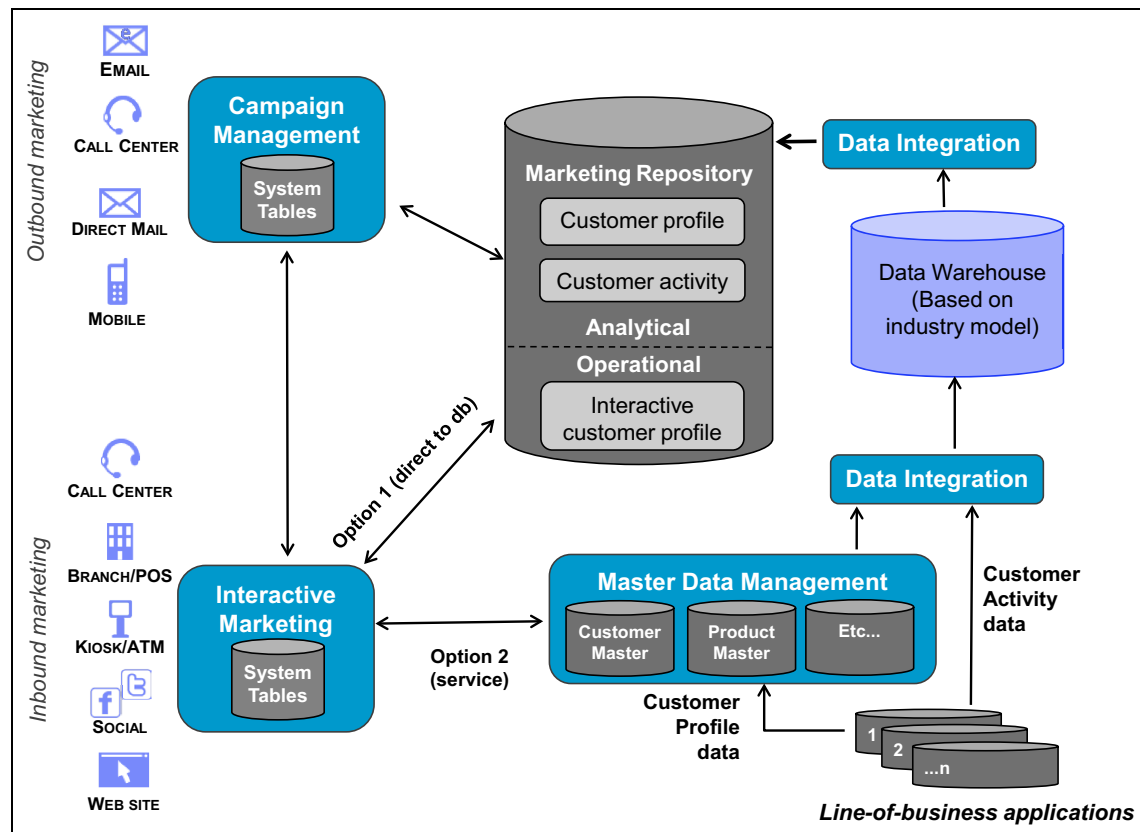
Figure 1-1   Solution overview

Information about customers and prospects comes from different applications and data sources, with some being internal applications and some being external data sources. Often the internal applications are specific to the line-of-business. The information from these applications includes both customer profile information (contact information, identifiers, some demographics) and also customer activity information (product purchases, account activity and details, website activity, and so on). The customer profile information flows into MDM, where it is matched and resolved into a single trusted, accurate view. The customer activity information flows into a data warehouse.

The data integration component manages this flow of information, both into MDM and into or out of the data warehouse and marketing repository. Data integration includes capabilities such as these examples:

► The ability to extract, transform, and load (ETL) data to and from various applications and databases.

► The ability to apply data quality checks and data quality cleansing and standardization functions.

► The ability to use and share common metadata across the whole information supply chain.

► A business user glossary of terms that all participants in the information supply chain can use for a common understanding of data elements and definitions.

These capabilities ensure that information can both be moved from point A to point B, and be governed along the way.

Next in the information supply chain is typically a data warehouse. In this book, we describe how the information can be organized in a data warehouse to align with an industry-specific data model. This industry-specific model is particularly valuable when managing the customer activity information. Although customer profile information tends to have the same model or structure across industries, the customer activity information tends to look quite different from one industry to the next.

The last stop on the information supply chain in Figure 1-1 on page 6 is the marketing repository. The distinction of the marketing repository from the data warehouse is that the marketing repository is purpose-built for marketing. The data warehouse might typically have a superset of information that is used for multiple business functions across the enterprise. The marketing repository is purpose-built for marketing, with data structures that are most efficient for managing marketing campaigns and data values that are summarized or calculated in ways that are useful for marketing. Figure 1-1 on page 6 depicts a single logical marketing repository with separate analytical and operational

sections. In a physical implementation, the analytical and operational sections should be physically separate for optimal performance.

After the customer profile and customer activity information are combined in the marketing repository for a complete, trusted 360-degree customer view, the campaign management and interactive marketing components leverage that trusted view to optimize both outbound and inbound marketing interactions. The interactive marketing component in Figure 1-1 on page 6 includes two options for retrieving customer information needed to make a real-time marketing decision.

► The first option, and most commonly used option, is to retrieve the necessary interactive marketing profile for a customer directly from a dedicated marketing repository table (or cache). This offers the lowest latency for retrieving information, which is critical when you are dealing with marketing interactions in websites or mobile applications, where every fraction of a second is apparent to the consumer.

► The second option is to retrieve the interactive marketing profile from an MDM service, or to retrieve part of the interactive marketing profile from MDM to augment what is in the marketing repository. This option might be used if there is more up-to-date information in MDM than in the marketing repository, and that up-to-date information is critical to making the marketing offer decision.

Because the second option might introduce more latency for making the service call, it might be more appropriate to use when the marketing interaction is through, as examples, a call center, kiosk, or automated teller machine (ATM), where the addition of an extra fraction of a second is not as noticeable to the consumer.

1.4 Mapping IBM software to the solution

The previous section (1.3, "Combining trusted information and Enterprise Marketing Management" on page 6) describes general components that make up the overall solution to combine trusted information and Enterprise Marketing Management. IBM has software products that are perfectly suited to fulfill the requirements of each component of the overall solution. Figure 1-2 is another depiction of the solution, this time with the appropriate IBM software product assigned to each component of the diagram.
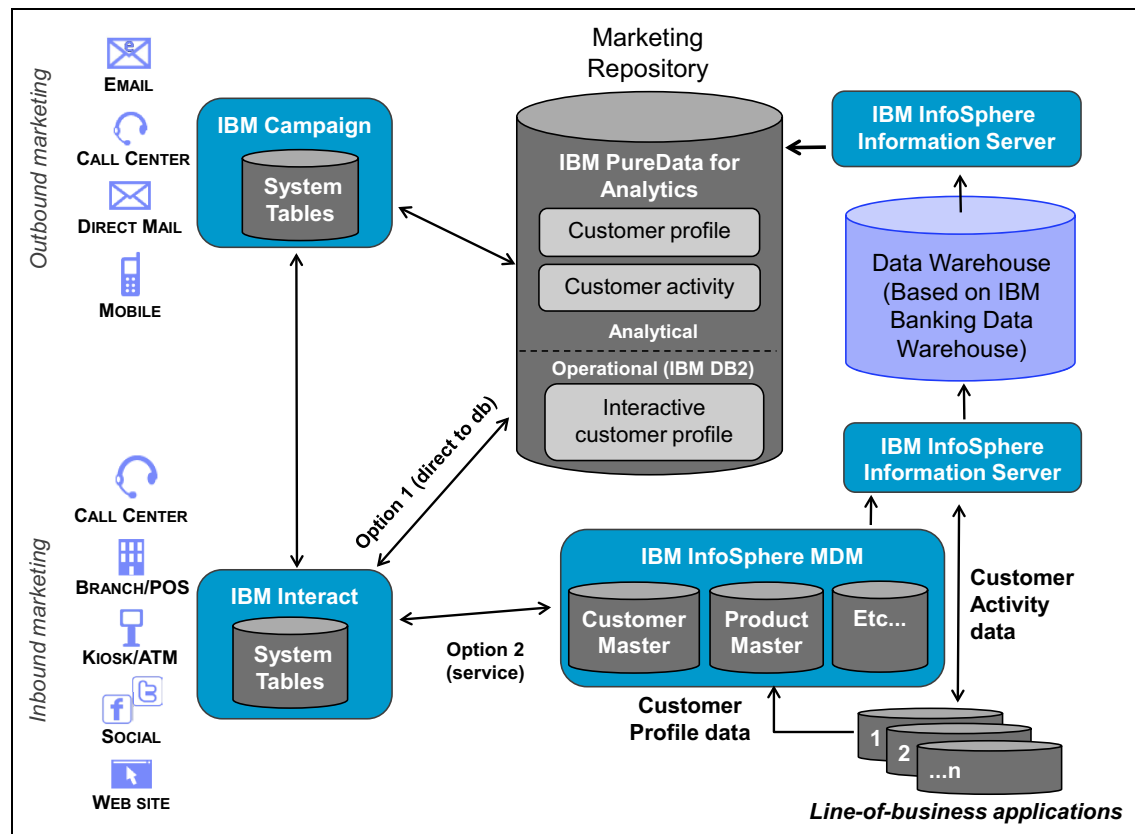
Figure 1-2   Solution Overview with IBM Software

The subsequent chapters of this book provide more detail about how each of these IBM software components is used to create the trusted customer information and act on that information. Each chapter uses a technical blueprint that depicts the solution components with more detail. The overall picture of that "blueprint" is shown in Figure 1-3.
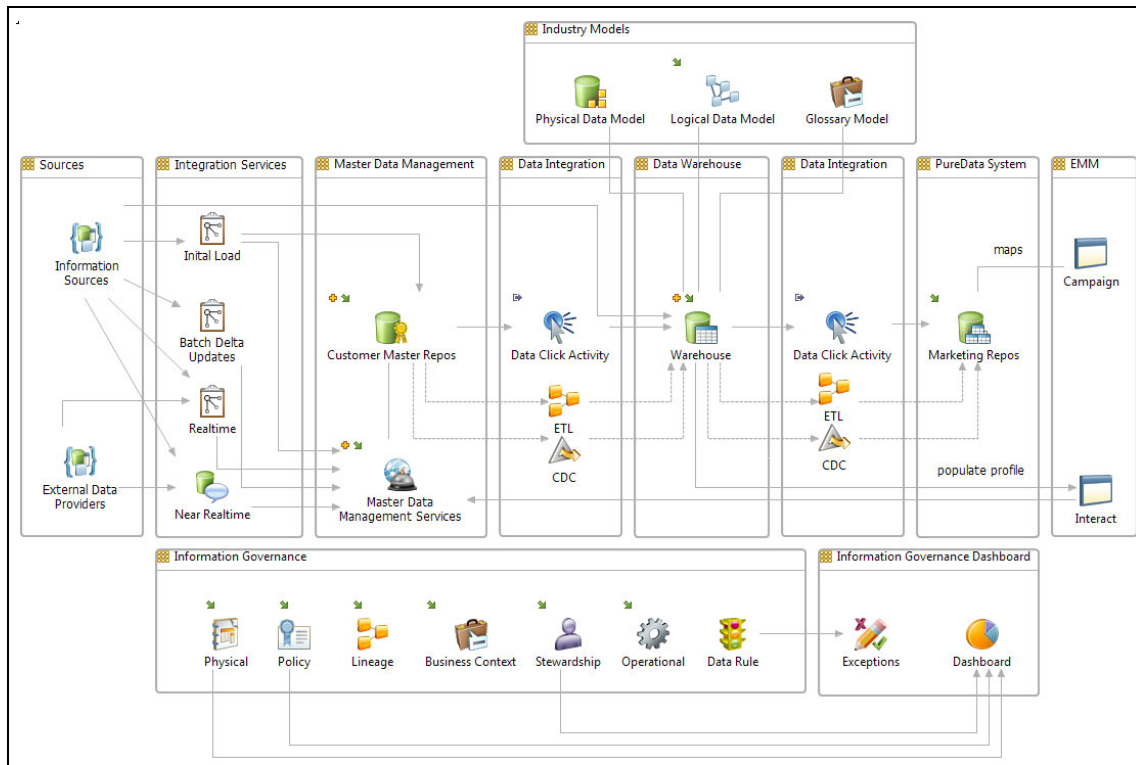
Figure 1-3   Solution technical blueprint

1.5 Example marketing scenarios

Throughout this book, we use an example to illustrate how each component supports two example scenarios: one outbound marketing scenario and one inbound marketing scenario.

Each example scenario is for a fictional financial services institution. This retail banking institution is planning a multi-channel *winback* campaign designed to target customers who are most likely to close their accounts or churn. These customers will be targeted with offers designed to increase their satisfaction with the bank's services and reduce the inclination to close their accounts. The campaign will include outbound offers that can be distributed through mail, email, or social media, and also inbound offers during a customer's interaction on the bank's website. The EMM platform manages the definitions of campaigns, offers, segmentation, and marketing rules. And the Information Management capabilities create the underlying trusted 360-degree customer view upon which the marketing decisions are based.

2

Using IBM Industry Models for marketing

In this chapter, we review two options for creating the marketing repository:

► Using a stand-alone data mart
► Using a data mart integrated with a data warehouse

When creating a data mart, the main steps are as follows:

1. Requirements analysis
2. Logical data modeling
3. Physical data modeling and database implementation
4. Data mart population

The two main sources of data for the marketing repository are customer data from the Customer Master Repository and product and account data from the source systems.

This chapter focuses on requirements analysis and logical data modeling. Strategies for populating the data mart are described in Chapter 4, "PureData for Analytics: The marketing data repository" on page 57. Physical data modeling is outside the scope of this document.

13

2.1 Two approaches to creating the marketing repository

The two main approaches to creating the marketing repository are as follows:

► Create a stand-alone data mart that contains only the fields required by IBM Campaign.

► Integrate the fields required by IBM Campaign into the Enterprise Data Warehouse (EDW) and then create a dependent marketing repository data mart with the fields required by IBM Campaign.

Both approaches provide you with the information needed in the marketing repository. This section describes the features and benefits of each approach.

2.1.1 Approach 1: Stand-alone data mart

The stand-alone data mart is quick to setup and can be designed so that the Customer Master Repository tables are directly copied from MDM into the data mart with no modifications. This reduces the modeling effort to the fields that are not sourced from the Customer Master Repository: typically account, transaction and product summary data. This approach can take advantage of IBM InfoSphere Data Click to move the customer data to the mart that is documented in 4.5, "Data Click" on page 73. An alternative data movement solution is required for the data from other data sources.

Advantages of this approach are as follows:

► Increased speed: No requirement to integrate with existing extract, transform, and load (ETL) jobs and data warehouse

► Reduced data modeling constraints: No requirement to integrate new data fields into an existing data warehouse model

Disadvantages of this approach are as follows:

► Data movement duplication: The required data might already exist in the data warehouse but must be extracted from the operational applications again.

► Data calculation duplication: Required calculations might be required for other purposes causing duplication and potential inconsistencies

► No reusability: Any fields used in the marketing repository are unlikely to be reused for other purposes.

► Governance requirements: Might need additional governance for new data movement flows and a new data mart.

2.1.2 Approach 2: Enterprise Data Warehouse (EDW)

In this scenario, all the fields required by IBM Campaign must be available in the EDW. When the IBM Campaign input data requirements are known, they are mapped to the EDW and the delta list of missing elements is created. These missing elements must be made available in the EDW before they can be included in the marketing repository data mart. Missing elements will typically either be sourced from the Customer Master Repository or from an operational application.

IBM Industry Models can be used to accelerate the creation of the EDW and can also be used to extend an existing EDW to support any new attributes.

Advantages of this approach are the following:

► The marketing repository only sources data from the EDW, simplifying data movement and avoiding data movement redundancy
► All new fields added to the EDW and available to be used for other business requirements
► All data fields are governed as part of the EDW

Disadvantages of this approach are the following:

► Adding fields to the enterprise data model and EDW may be more complex than creating a stand-alone data mart

The remainder of this chapter focuses on approach 2, and assumes the use of IBM Industry Models to extend the EDW before creating the marketing repository.

2.2 Using IBM Industry Models to design the marketing repository

In this section, we review the solutions available from IBM to accelerate the design and implementation of the EDW and the marketing repository.

2.2.1 IBM Industry Models value proposition

IBM Industry Models combine deep expertise and industry best practice in a set of extensible blueprints for use both by business and IT communities. The models are used to accelerate the creation of an industry-specific Enterprise Data Warehouse and also data marts for business intelligence and reporting requirements. As part of the IBM Information Management portfolio, the industry

models are based on the experience of working with more than 500 clients, and more than fifteen years of development work.

Each comprehensive industry model contains the following items:

- ► Business terminology components to define and refine business requirements that improve communication between business and IT
- ► Business intelligence analysis templates that accelerate the collection of business reporting requirements for a top-down modeling approach
- ► Data warehouse design components to accelerate building out the database structures that support the business requirements
- ► Mappings between the business terminology, analysis templates, and design components that accelerate transforming business requirements into database designs

These proven industry model components provide the following benefits:

- ► Industry-specific content covering a wide range of standard industry business concepts and practices and also coverage of major regulatory requirements
- ► Extensive content and documented usage methodology to reduce project risk
- ► Project acceleration through the use of prebuilt models to reduce data modeling effort

2.2.2 Using IBM Industry Models to accelerate implementation

IBM has a range of industry-specific and cross-industry data model offerings that can be used to accelerate the design and implementation of a data warehouse and down-stream data marts, including the marketing repository. IBM Industry Models offerings cover the following industries:

- ► Banking and financial markets
- ► Healthcare: provider and plan
- ► Insurance
- ► Retail
- ► Telecommunications

The full list of IBM data model offerings can be found at the following URL:

http://www.ibm.com/software/data/industry-models/

When using an IBM industry model, you start by identifying the prebuilt model components that support your data and business requirements. Rather than modeling every concept from the beginning, use the model content as a starting point to support your standard requirements. You can then focus your data

modeling efforts on the customizations that make your organization's business unique.

The remainder of this chapter is focused on the process for gathering data modeling business requirements, extending the design of the EDW using the Banking Data Warehouse (BDW) and creating a marketing repository data mart.

2.3 Requirements for the marketing repository

The marketing repository is the data mart that will contain all the input data required by IBM Campaign. It requires customer and account data, which can be divided into the following subject areas:

► Core customer information
► Customer demographics
► Contact preferences
► Customer accounts
► Account transactions
► Customer scores
► Product and product usage summaries

This type of data is required by many lines of business within the organization for a wide variety of analytical purposes. With the opportunities for reuse, the data should be integrated into and sourced from the organization's EDW to provide the necessary data lineage, governance, and quality for the data that will be used to generate the marketing campaigns and offers.

If an EDW is not used as the source for the marketing repository data, a separate set of data feeds will be required to populate the marketing repository directly from the source systems adding to the proliferation of disparate data stores and point to point ETL jobs.

2.4 Data requirements analysis

The purpose of data analysis is to gain a thorough understanding of the data elements that are required for the marketing repository and from what systems they should be sourced. If you already have an EDW, then you should map your required data elements to the data warehouse and identify any gaps. Each data element represents a required field in the EDW, the marketing repository, or both.

2.4.1 Review and prepare the data requirements

Data requirements can be provided to you in many formats including functional requirements documents, layout diagrams, or spreadsheets of data elements. In whatever format you receive your requirements, document them with at least the following details for each data element:

► Data element name

  The data element name is the name with which the business users are familiar. The name should be as unique and as descriptive as possible.

► Data element description

  The data element description must clarify any limits or restrictions on the data element. For example, data elements relating to customers should specify whether all customers are in scope or if just retail or corporate customers are being used.

► Data element source

  The data element source describes from where this particular field needs to be sourced. Is the data element already in the data warehouse, is it in the MDM solution or does it need to be sourced from an operational application?

After the required data elements are documented, the elements that need to be modeled as part of the EDW can be identified.

2.4.2 IBM InfoSphere Business Glossary for business requirements

One method for documenting business requirements is to use IBM InfoSphere Business Glossary. The IBM InfoSphere Business Glossary can be populated by creating a set of business terms or by importing the business terms that are included with the Banking Data Warehouse. Each business term represents a data element or requirement.

You will need an IBM InfoSphere Business Glossary Category that you can use to gather your business requirements. The Banking Data Warehouse methodology recommends that this category be created within the *Scopes* category. You also need a custom attribute for business terms referred to as Data Element Source with the attributes as shown in Figure 2-1 on page 19.

Figure 2-1   Create New Custom Attribute

You also need the authority to edit the new IBM InfoSphere Business Glossary Category and create business terms.

Edit the IBM InfoSphere Business Glossary Category and choose the individual business terms that describe your project's data element requirements. To reuse existing IBM InfoSphere Business Glossary terms from the Banking Data Warehouse, start entering the business term name in the References Business Terms field. IBM InfoSphere Business Glossary then suggests a set of business terms based on what you type. Click the term you want to use. That term you select is then listed below the entry field. Do this for each term you want to include in your category and then save your work.

If you need to include business terms that are not already in the glossary, create each term in its appropriate category based on what the business term describes. The Banking Data Warehouse provides a full category hierarchy to contain all its business terms that can then be reused in other categories. Any new business term that will not be reused elsewhere can be added to this category by using the Contains Business Terms field.

We use the References Business Terms field to access existing business terms without moving them from their original location; this ensures that business terms are reused and still available in other categories for other purposes.

IBM InfoSphere Business Glossary Category containing a set of project requirements is shown in Figure 2-2. The terms sourced from the Banking Data Warehouse business terms are *referenced business terms* and new terms that were created are *contained business terms*.

Figure 2-2   Project requirements in IBM InfoSphere Business Glossary

After you include all the business terms in the category, edit each one and complete the Data Element Source field that was created earlier (see the example in Figure 2-3). If you are sourcing your marketing repository data from an EDW, use `Data Warehouse` as the entry in the Data Element Source field for any data requirements that are already in the data warehouse. Then, identify all

the fields that exist in the Customer Master Repository, but not the data warehouse, by using MDM as the entry in the Data Element Source field. Any elements that come from other data sources should use `Source System` as the entry in the Data Element Source field.

If you are not using a data warehouse, then identify all elements that exist in the Customer Master Repository first and then elements that exist in the source systems.

After you capture all your data element requirements in the category, a report in Metadata Workbench produces a listing of your requirements that can be exported to a spreadsheet for use by other team members. An example of the report is shown in Figure 2-3.

References Terms - Name	References Terms - Short Description	References Terms - Data Element Source
Address name	The name of a building or organization which appears in an address e.g. Rose Cottage, Bramble Way or General Post Office, Dublin 2	Data Warehouse
City	Identifies a Geographic Area that is a municipality with legal power granted by a state charter.	Data Warehouse
Country	Identifies a Geographic Area that is recognized as an independent political unit in world affairs.	Data Warehouse
Date of birth	Identifies the birth date of the Individual.	Data Warehouse
Email address	Identifies an Electronic Address which identifies a logical address which can be used to send and receive correspondence over a computer network; for example, jsmith@jonesinc.com.	MDM
Gender	A dimension that classifies measures according to the sex of Individuals.	Data Warehouse
Given name	Identifies an Involved Party Name Component that is the part of a name that is chosen and assigned to an Individual at his or her birth, usually by the parent(s); for example, John, Jurgen, Helga or Jane.	Data Warehouse
Individual certification type	A Classification Scheme that distinguishes between Individuals according to the highest level of official accreditation awarded for their skills.	MDM

Figure 2-3   Category Summary Report from Metadata Workbench

You now know all the data elements that are required for the marketing repository data mart and whether the required fields are already in the EDW or they need to be brought into the EDW.

After the data requirements documentation is complete, the project team business analysts should review it for accuracy and omissions.

2.5 Use the BDW to model your data requirements

Your completed and validated data requirements are the input to the data modeling process. Complete the following two data modeling exercises (described in this section):

1. Extend the EDW logical model to support any new attributes that are required for the marketing repository that are not already in the EDW.

2. Create the marketing repository Data Model.

2.5.1 Extend the EDW logical model

The purpose of a logical data model is to support the integration of information from multiple source systems and to organize this information by identifying a single location for each unique piece of business information. A logical model is not aware of or sensitive to the underlying DBMS where it might be eventually deployed.

If you have an existing EDW model, you must extend it to accommodate the new data element requirements. These new requirements can be identified by sorting your list of data element requirements from 2.5.2, "Create the marketing repository logical data model" on page 24 to highlight where the Data Element Source field shows either `MDM` or `Source System`.

If you already have an EDW based on the BDW, it might look similar to Figure 2-4 on page 23. Entities of the same color are part of the same data model subject area or package.
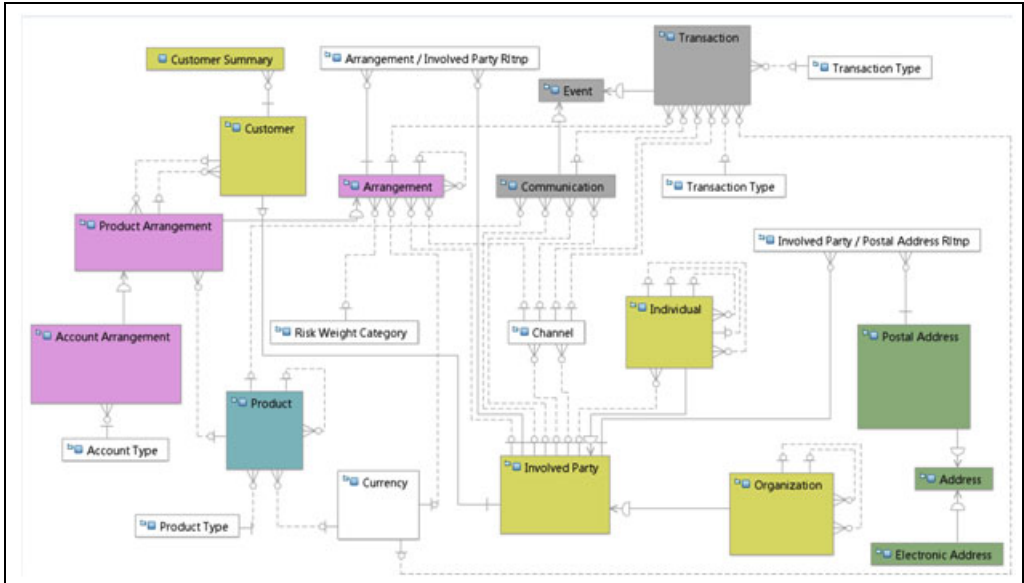
Figure 2-4   Banking Data Warehouse logical model

Table 2-1 lists the common types of data elements that are used for IBM Campaign and where they can be modeled using the BDW.

Table 2-1   Common data elements for IBM Campaign

Data element type	BDW entity
Customer Name	Individual
Customer Address	Postal Address
Email and Phone Numbers	Electronic Address
Customer Demographics	Individual
Account	Arrangement
Account Transaction	Transaction

Using the table as a guide, add the new data elements to your logical data model. For some data elements that do not already exist in the BDW model, such as contact preference opt-out flags, you might need to create a new entity and attributes.

Any new fields added to the EDW logical model must be shared with the ETL development team so the fields can be populated with the correct source data.

2.5.2 Create the marketing repository logical data model

When all the data elements exist in the EDW, the marketing repository data mart can be modeled. Typically, a data mart uses more denormalized data structures than the EDW so some entities in the EDW might be collapsed into a single entity in the mart. The BDW dimensional model contains many useful structures for designing data marts.

The design goal is to model a data mart that is consumable by IBM Campaign and is easy to maintain and manage.

An example of a logical model for the marketing repository is shown in Figure 2-5. The entities in blue represent content that is typically sourced from the Customer Master Repository; the entities in yellow show content that might be sourced from the Customer Master or other data repositories.
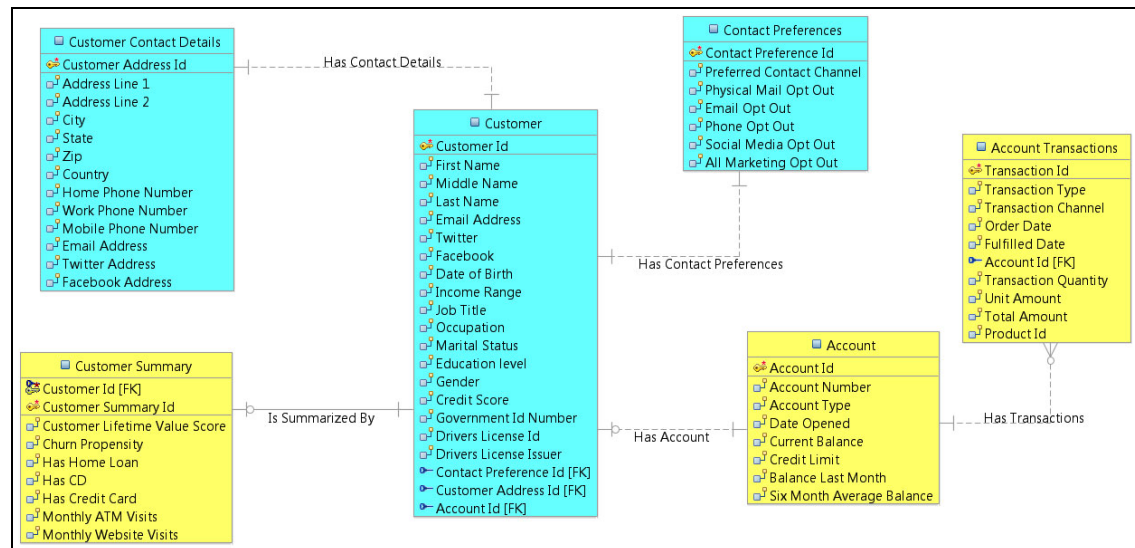
Figure 2-5   Sample marketing repository Data Model

Using InfoSphere Data Architect, creating dependencies between entities and attributes within and across models is possible. For example, you can show that an attribute in the marketing repository Data Model is derived from or sourced from an attribute in the EDW Data Model. This can be used to show sourcing relationships such as "is DW source for" to provide greater context to the models and to ensure better impact analysis and governance.

With the IBM InfoSphere Business Glossary Eclipse plug-in installed, InfoSphere Data Architect allows you to assign business terms to entities and attributes in a

logical model and propagate those business term relationships to the physical model. This capability provides data modelers with useful business context during their modeling activities and builds rich metadata.

2.5.3 Create the physical models and Data Definition Language

After you complete the logical models, you can transform them into physical data models that are platform-specific by using InfoSphere Data Architect. If you select the **Generate traceability** option when you transform the models, you automatically create metadata mappings between the logical and physical models that provide traceability.

A database administrator (DBA) can review the physical model and customize it to ensure appropriate performance on the selected database platform. After the physical data model is complete, InfoSphere Data Architect can generate the Data Definition language (DDL) that the DBA will use to create the database tables.

If you are creating your EDW for the first time or are adding customer information to the EDW or data mart for the first time, you can use IBM InfoSphere Data Click to move the Customer Master Repository tables directly into the EDW or data mart; that means you can skip the steps for creating a table and loading a table.

If your target is the data warehouse, the layout of the source tables can be restructured to more closely align to a typical BDW model by creating a view across the Customer Master Repository tables before you use IBM InfoSphere Data Click to make the tables look like those created from BDW. An example of several views is shown in Figure 2-6.

MDM db table	MDM table column	BDW Entity Name	BDW Attribute Name	MDM Views to map to BDW
CONTEQUIV	ADMIN_SYS_TP_CD	Involved Party	Source System Id	View 1- Join to make Involved Party
CDIDTP	ID_TP_CD	Involved Party	Unique Id In Source System	View 1- Join to make Involved Party
PERSONNAME	GIVEN_NAME_ONE	Individual	Given Name	View 2- Join to make Individual
PERSONNAME	GIVEN_NAME_TWO	Individual	Middle Name	View 2- Join to make Individual
PERSONNAME	LAST_NAME	Individual	Surname1	View 2- Join to make Individual
CONTEQUIV	ADMIN_CLIENT_ID	Individual	Social Security Number	View 2- Join to make Individual
PERSON	BIRTH_DT	Individual	Birth Date	View 2- Join to make Individual

Figure 2-6   Example table views

2.6 Export your logical and physical models to Metadata Workbench

After you creating your logical and physical models, use InfoSphere Metadata Asset Manager to import the models into Metadata Workbench. You can also import the database metadata that was created when you instantiated the physical model. This makes the tables and models available in the Information Server repository for the ETL team and also for governance and management purposes.

When you import a physical data model with InfoSphere Metadata Asset Manager, select the **Import LDM assets with DBM** option. This automatically imports the logical model that was used to create the physical model and all the traceability mappings between the two models in one step. For this type of import to run successfully, the logical and physical models must be in the same folder on the source machine. In Metadata Workbench, the traceability mappings are displayed as `Implements Logical Entity` or `Implemented by` relationships, automatically creating valuable metadata.

After the logical and physical models are imported into Metadata Workbench, you can export business term assignments to entities, attributes, tables, and columns from InfoSphere Data Architect to the repository. Now, when you click a business term in IBM InfoSphere Business Glossary or Metadata Workbench, you see the logical and physical model components that relate to each business term in its Assigned Assets section.

2.7 Chapter summary

This chapter describes the two approaches to building the marketing repository:

► Creating a stand-alone mart
► Sourcing all the data from an EDW

The chapter listed the advantages and disadvantages of each. It also reviewed requirements analysis using IBM InfoSphere Business Glossary and logical model design using BDW.

3

Configuring MDM with marketing-relevant attributes

Configuring InfoSphere Master Data Management (MDM) for customer and prospect data to take account of marketing-relevant attributes allows you to drive marketing campaigns. Campaigns based on this reliable information about customers and prospects means the right offer can be made to the right person. Creating the right offer using Enterprise Marketing Management (EMM) software is described in Chapter 6, "Leveraging Customer Data within IBM Campaign" on page 103 and Chapter 7, "Leveraging Customer Data within IBM Interact" on page 127.

This chapter describes the value that a customer master brings, how to install InfoSphere MDM, how to populate a person entity in MDM with marketing-relevant attributes, deployment, and testing.

When person-data is in MDM, MDM processes will run to ensure that the data is trustworthy. The information supply chain takes the MDM person-data and moves it through to the marketing repository, where it is used by EMM to target offers to the right person.

For more information, see the following sources:

► IBM InfoSphere Master Data Management products:

  http://www.ibm.com/software/data/master-data-management/

► EMM:

  http://www-03.ibm.com/software/products/en/category/SWX00

EMM contains several products. The use cases considered in this book relate to the following items:

► Outbound marketing using IBM Campaign:

  http://www-03.ibm.com/software/products/en/campaign-management/

► Real-time inbound marketing using IBM Interact:

  http://www-03.ibm.com/software/products/en/real-time-inbound-marketing/

3.1 The value of a customer master

Often data about the people your enterprise knows about is locked in silos, such as proprietary databases that are accessible only by proprietary applications. For example, person information might be stored in a database for the accounts department, and accessed by an accounting package. This information about people, including customer and prospect, is hidden from the other parts of the enterprise. Many silos that contain information about people might exist across the enterprise, and the data in these silos about customers and prospects can be inconsistent, incomplete, missing, duplicated, or just plain wrong. If your person data is like this, how confident are you that decisions you make and the marketing campaigns you do based on this information will be successful?

Introducing master data into a company means that you get a single trusted view of customers and prospects. How do you know it is trusted? Because MDM has processes specifically to ensure this. The remainder of this section describes a customer master and how to implement it with IBM InfoSphere MDM. Your enterprise can then be ready to provide a solid foundation of trusted information to build your marketing campaigns on. This chapter covers installation, configuration, deployment and testing.

Styles of master data

MDM is used to build a trusted view of the customer. MDM has three main styles:

► Registry style is a virtual registry view that provides a trusted view across many sources of customer information.

► Centralized style is a central repository, where customer information is copied and mastered (authored).

► Hybrid style is a combination of the registry and centralized styles.

How do you decide which style is appropriate for you? A simple way of getting the benefits of master data, without changing the existing data sources, is to use the registry style (providing a virtual registry view), and then have the consumers of customer information take that information from this new trusted view.

To author the master data centrally, giving you complete control over its lifecycle and quality, use the centralized MDM style, which meets this need.

To use parts of each style, then the hybrid style is appropriate.

IBM implements the MDM capability with the IBM InfoSphere MDM product, using the following terminology:

► Virtual MDM implements the registry style.
► Physical MDM implements the centralized style.
► Hybrid MDM implements both registry and centralized styles.

A useful resource is the MDM Developers community:

http://ibm.co/mdmgroup

This book details an information supply chain to feed an EMM marketing solution. A vital information source is the trusted person data in physical MDM. An advantage of using a physical MDM approach with EMM solutions is that there is only one repository (Customer Master Repository) that the information integration processes must be aware of to access customer and prospect details.

3.1.1 Person identifiers

A person can be identified in many different ways, such as from the person's name and address, Social Security number, and telephone number, and through a social network or by employee number. Different systems might use different identifiers for the same person. Customer master data can include any number, and all types, of identifiers associated with a person. This also means that consumers of customer master data can get person information using the identifier with which they are familiar. Managing all identifiers that are associated with a person is a major strength of MDM.

MDM is concerned with the following two main types of person identifiers:

- ► Business keys, including name, address, and unique identifiers such as Social Security number, are all customizable. These keys are used when matching parties to see whether records are duplicates.

- ► Contact equivalences are identifiers that another system uses to identify parties are also customizable.

3.1.2 Increasing confidence in your customer data

Various processes in MDM can be used to increase the confidence in your customer and prospect data.

A big issue with customer information is duplicates; deduplication and survivorship rules address this issue. In addition, MDM can normalize and validate information.

Deduplication

Certain identifiers associated with a person may seem unique (such as social security number) and others less so (such as name). Even identifiers that should be unique, could have been input incorrectly. MDM has processes to recognize duplicate entries and merge them if appropriate. This merging (or collapsing) is known as *deduplication*.

Survivorship rules

IBM InfoSphere MDM has the Probabilistic Matching Engine (PME), which is used to match person data. It uses a statistical approach to score whether two person records match. If there is a close, but not perfect, match between two people, each person record is allocated a score and corresponding suspect categorization. When there is a guaranteed match, the information might simply be merged to form a single, golden record.

You can customize survivorship rules to determine exactly what you want to do when either of the following actions occur:

- ► A new person is added to the MDM operational hub and there is an existing person.

- ► A data steward collapses records or during *evergreening* (where person data is added into the hub then deduplicated and cleansed later).

Survivorship rules allow you to specify what data from the source person records get collapsed into the merged record,

3.2 Open Service Gateway initiative (OSGi)

OSGi reduces complexity by providing a modular architecture for enterprise applications. IBM InfoSphere MDM V11 uses OSGi to manage changes and customizations. In previous versions of MDM, changes were merged with the MDM system code to create a new enterprise archive (EAR) file, which was then deployed. In V11, customizations are managed and deployed separately from the MDM system code, (so there is no need to deploy the MDM system code when deploying your changes) in Composite Bundle Archives (CBAs). The MDM Workbench uses a mode-driven tooling approach, which allows customizations to MDM to be created and deployed as CBAs.

For more information about OSGi, see the OSGi Alliance website:

http://www.osgi.org/Main/HomePage

3.3 Business Process Management (BPM)

MDM tasks, such as data stewardship or approvals, often require a workflow. With BPM, you can create user interfaces and workflows. Using the MDM Application Toolkit, you can create BPM user interfaces that work on your master data.

3.4 IBM InfoSphere MDM V11 installation on Windows

When you consider how to install InfoSphere MDM V11, you have several decisions to make:

► Which edition?

– Standard Edition gives you virtual MDM.
– Advanced Edition gives you virtual, physical, and hybrid MDM.
– Enterprise Edition additionally gives you a collaborative authoring capability.

► Custom or typical installation?

– Custom installation allows you to use other supported software and is more flexible.

– Typical installation provides software set that allows you to run the MDM stack.

► Server or workstation installation?

   – Server installation gives you only the server.
   – Workstation installs the server and a workbench environment.

This chapter describes the typical MDM Advanced Edition, workstation installation on Windows.

3.4.1 Planning the installation

When you plan the installation, complete the following tasks:

► Check the system requirements of your machine. Consider having at least 150 Gb of free hard drive space and 8 Gb of RAM prior to installation.

► Uninstall IBM WebSphere® Application Server from your machine and delete all WebSphere profile folders.

► Uninstall all IBM DB2® software. Delete user database folders. The DB2 folder needs to be deleted and also the GSK (GSK is the name of the folders in the program files directories) folders. Also, you must remove the DB2 administrator user ID (db2admin by default).

► Download the required software. You should have the following files:

   – `DB2_ESE_10_Win_x86-64.exe`
   – `IMDM_STND_ADV_ED_V11.0_P1.tar.gz`
   – `IMDM_STND_ADV_ED_V11.0_P2.tar.gz`
   – `IMDM_STND_ADV_ED_V11.0_P3.tar.gz`
   – `IMDM_STND_ADV_ED_V11.0_P4.tar.gz`
   – `IMDM_STND_ADV_ED_V11.0_P5.tar.gz`
   – `IMDM_WKBCH_STND_ADV_ED_V11.0_P1.tar.gz`
   – `RAD_WS_8.5.1_1.zip`
   – `IMDM_WKBCH_STND_ADV_ED_V11.0_P2.tar.gz`
   – `IMDM_WKBCH_STND_ADV_ED_V11.0_P3.tar.gz`
   – `INSTL_MGR_1.6_RQ_INSTL_MDM11_WIN.zip`
   – `INSTL_STUP_TK_IMDM_V11.0.tar.gz`
   – `RAD_WS_8.5.1_2.zip`
   – `RAD_WS_8.5.1_3.zip`
   – `RAD_WS_8.5.1_4.zip`
   – `WAS_V8.5_1_OF_3.zip`
   – `WAS_V8.5_2_OF_3.zip`
   – `WAS_V8.5_3_OF_3.zip`
   – `v10.1fp2_ntx64_server.exe`
   – `DB2_ESE10.1_WIN_IMDM_V11.0_P2.tar.gz`
   – `DB2_ESE10.1_WIN_IMDM_V11.0_P1.tar.gz`

► Download WebSphere Application Server V8.5 Fix Pack2 from the following web page:

http://www.ibm.com/support/docview.wss?uid=swg24034672

You need the following files:

– 8.5.0-WS-WAS-FP0000002-part1.zip
– 8.5.0-WS-WAS-FP0000002-part2.zip

A valid license for IBM Rational® Application Developer for WebSphere (which we will refer to as RAD) is required. Complete the following installation actions:

1. Install an application that can expand the tar.gz files on Windows.

2. Put the expanded files in a folder structure, such as the one described in Table 3-1.

Table 3-1   Folder structure and expanded files

Folder	Description
DB2	For the IBM DB2 software
DBFP	For the IBM DB2 fix pack
MDM	For the MDM files
MDMWB	For the MDM Workbench files
RAD	For Rational Application Developer (Be sure you do not create a second RAD folder under this one.)
WAS	For the WebSphere Application Server V8.5 files
WASFP	For the WebSphere Application Server V8.5 fix pack.

3.4.2 Installing InfoSphere Master Data Management

For the expanded files in 3.4.1, "Planning the installation" on page 32, complete the following steps:

1. Navigate to `MDM/disk1` location and run the **`launchpad64.exe`** command as administrator.

   Figure 3-1 shows the initial launchpad window.
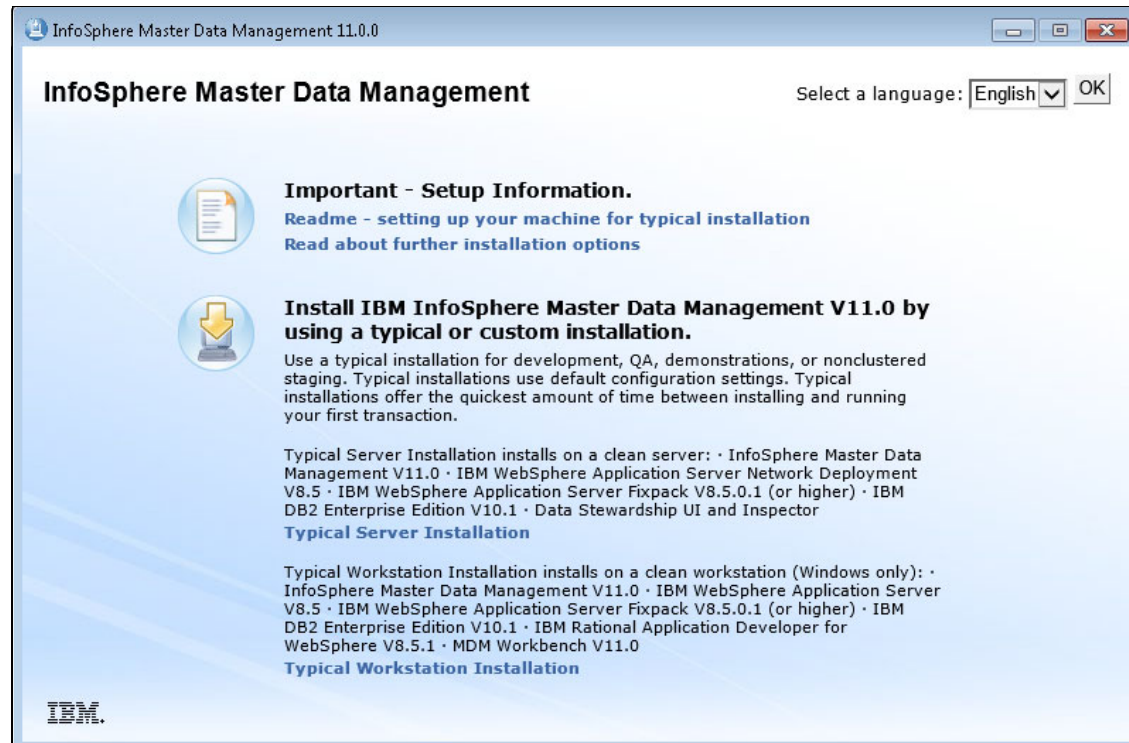
Figure 3-1   Launchpad installation options

2. Click **Typical Workstation Installation**, and then follow the windows instructions, accepting the defaults until you see the final Install page, as shown in Figure 3-2 on page 35.

3. Click **install**.
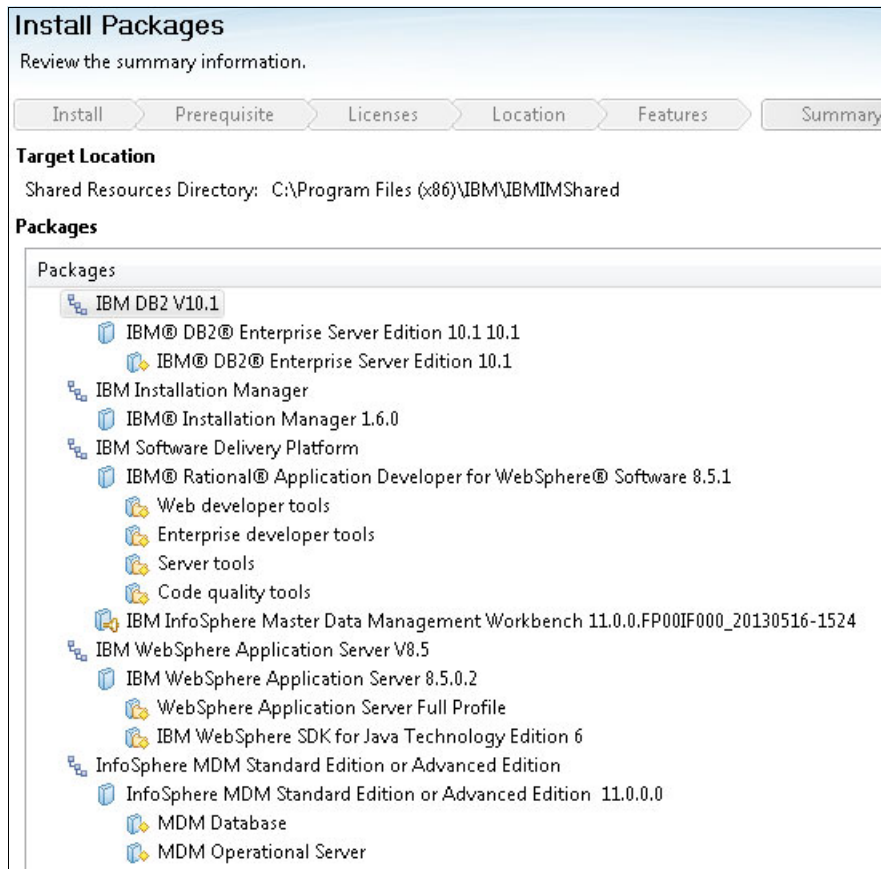
Figure 3-2   Final Install page

When the installation is complete, do the following steps:

1. Check that the installation verification tests (IVT) completed successfully. Do this by navigating to the following folder and opening the three files there:

   `C:\ProgramFiles\IBM\MDM\IVT\testCases\xml\response\`

   If the IVT was successful, the three files in this folder will all have XML content, which contains the following string:

   `"<ResultCode>SUCCESS</ResultCode>"`

2. Upgrade the trial license to a full RAD license. To do this, start Rational Application Developer and click **Manage License**. The Manage Licenses window opens, as shown in Figure 3-3 on page 36.

One way to add a license is to import the activation kit. See the Rational Application Developer documentation for more details. After you add your license, verify it in the Manage Licenses window.
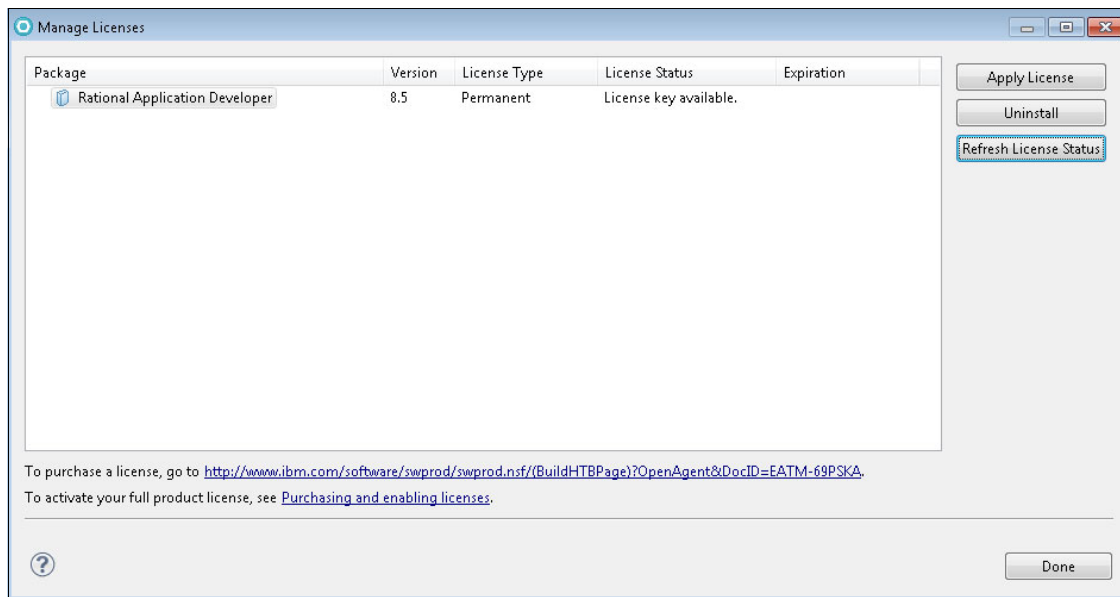
Figure 3-3   Rational Application Developer Manage Licenses

Some useful resources for working with the MDM Workbench are at the following web page:

http://ibm.co/PMjK9U

3.5 Adding marketing attributes to MDM

In this book, we assume that the marketer needs access to information about customers and prospects. The marketer has a list of person characteristics that are needed in the marketing repository. The customer details must be trusted so that the offers that are sent as part of the marketing campaigns are sent to the correct person. The customer master has a coherent single view of the person information.

The representative attribute types that a marketer might want for person details are indicated in the following list. The list is not exhaustive or mandatory.

► Source system code (identifier of a source system where person is known)
► Source system record ID (identifier of the person on the source system)
► First name
► Middle name
► Last name
► Address Line 1
► Address Line 2
► City
► State/Province
► Post code
► Country
► Social Insurance ID number
► Birth date
► Income
► Gender
► Language code
► Marital Status
► Education
► Home phone
► Email address
► Mobile phone
► Mail Opt In
► Phone Opt In
► Email Opt In
► Opt Out of All
► Preferred channel
► Preferred channel ID
► Income range code
► Credit Score
► Job title
► Job Title ID

The next section describes how to add this person information into master data.

Mapping marketer needs into the MDM Party model

Before adding values of attributes, such as the job title of a person, you must know where in the MDM data model each attribute should be stored.

This section describes how to take the attributes that the marketer needs and understand where they should be stored in the physical InfoSphere MDM Party model. Consider the following three ways for storing these attributes in MDM:

► Identify an existing field in the Party model that can be used to store the person attribute.

Table 3-2 lists attribute names, which the marketer supplied, the equivalent MDM business object and the MDM attribute name that already exist in the Party domain:

Table 3-2   Attribute to MDM attribute mapping

Attribute name	MDM business object	MDM attribute
Source system code	TCRMAdminContEquiv	Type
Source system record ID	TCRMAdminContEquiv	Num[a]
First name	PersonName	givenname1
Middle name	PersonName	givenname2
Last name	PersonName	LastName
Address Line 1	TCRMAddress	AddressLine1
Address Line 2	TCRMAddress	AddressLine2
City	TCRMAddress	City
State/Province	TCRMAddress	ProvinceStatetype
Post code	TCRMAddress	ZipPostalCode
Country	TCRMAddress	CountryCode
Social Insurance ID number	TCRMIdentifier	type 1[b]
Birthdate	TCRMPerson	Birthdate
Income	TCRMIncomeSourceBObj	annualAmount
Gender	TCRMPerson	Gender
language code	TCRMPerson	PreferredLanguageType
Marital Status	TCRMPerson	MaritalStatusType
Education	TCRMPerson	HighestEducationType

Attribute name	MDM business object	MDM attribute
Home phone	TCRMContactMethod	ReferenceNumber
Email address	TCRMContactMethod	ReferenceNumber
Mobile phone	TCRMContactMethod	ReferenceNumber
Mail Opt In	TCRMPartyContact-Method	SolicitationIndicator
Phone Opt In	TCRMPartyContact-Method	SolicitationIndicator
Email Opt In	TCRMPartyContact-Method	SolicitationIndicator
Opt Out of All	TCRMParty	SolicitationIndicator
Preferred channel	TCRMPartyContact-Method	preferredContactMethodIndicator
Preferred channel ID	ContactMethodStatusType	preferredContactMethodIndicator[c]

a. In the database this is ADMIN_CLIENT_ID in CONTEQUIIV
b. In the database this is CDIDTP table which has sample content. ID_TP_CD = 1 is for social security number. We suggest considering using a different number in this code table for different social security provider.
c. In the database for Preferred channel ID we suggest using METHOD_ST_TP_CD in CDMETHODSTATUSTP

See the MDM V11 information center to learn about the structure of Party:

http://pic.dhe.ibm.com/infocenter/mdm/v11r0/index.jsp

► Identify an existing code table to which to add new values.

Some attributes can be put into existing MDM code tables, such as in Table 3-3.

Table 3-3   New values for code tables

Attribute name	Code table	Comment
Id Issuer	CDIDTP	new entry for each social security issuer
income range code	CDCLIENTPOTENTP	new code table values for each range

The list in Table 3-3 is not exhaustive. You might need to add entries such as these:

– Entries into CDIDTP to add any other identifiers that you have (for example, driving license).

– You can use the TCRMContactMethod table to add social media URLs.

▶ Create an extension to Person to hold any other information. As examples, there may be attributes that do not fit into the existing categories. For example credit score and job title are not directly mapped to the model. However, the Physical MDM model is flexible and can be extended.

For example, for job title we introduce a new code table. We add the job title code type and the credit score into an Entity Type extension of TCRMPerson.

Using the MDM Workbench to define customizations

Now that you know where in the MDM data model the person attributes will be stored, you must extend MDM to be able to hold attributes that are not in the standard data model.

This section explains how to use the MDM Workbench to add a new code table to hold job title values and also create an extension to the person data model to hold the job title and credit score.

Task to use the MDM Workbench to extend Person

Start Rational Application Developer. The first time you do this it looks for WebSphere runtimes and displays the window shown in Figure 3-4.
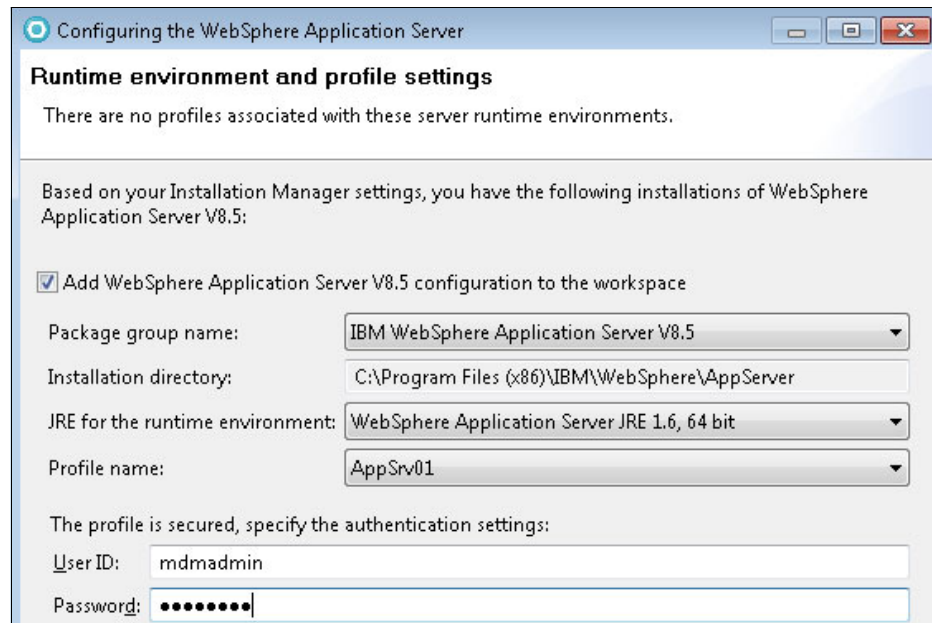
Figure 3-4   Profile settings

Enter the WebSphere user ID and password (this is `mdmadmin` for both by default, but you should change it for production systems).

The MDM Workbench uses a model-driven development approach. The model you work with to add customizations is the Development Module Model, which is stored in an MDM Development project. Using an intuitive Eclipse editor, customize the model and then generate the artifacts from it. This development project and the generated artifacts within it are included in an OSGi Composite Bundle Archive, which you deploy to the WebSphere Application Server. Complete these steps:

1. Create a new development project. Start by selecting **New** → **Development Project**, as shown in Figure 3-5. This project will hold the Person extension.
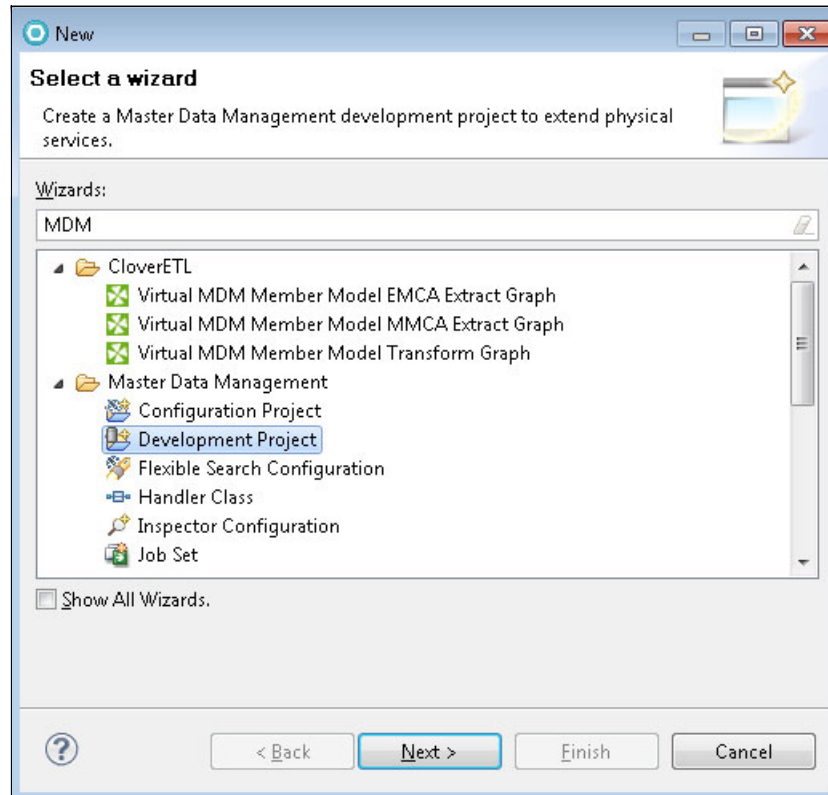
Figure 3-5   New MDM Development Project

2. When the new Project wizard starts, specify the following values:

    – Project name: `MarketingAttributeExtensions`
    – Base Java package name: `emmEnablement`
    – Namespace URL: `http://emmEnablement`

3. In the New Master Data Management Development Project, Master Data Management Development Project Settings (Figure 3-6), specify these values and then click **Finish**:

   – Identifier: `marketingEnhancements`
   – Database schema name: `db2admin`
   – Composite project: `marketingenhancements.cba`
   – Select the **Add development project to composite** check box.

Figure 3-6   Master Data Management Development Project settings

Open the `module.mdmxmi` file, which starts the Development Module Editor. In the editor do the following steps:

1. Create a new folder named `MarketingAttributeExtensions` by right-clicking the project and choosing new folder. This is an optional way to group artifacts in the model.

2. Create a new code table named `JobTitle` by right-clicking the folder and choosing new Code Table (Figure 3-7).
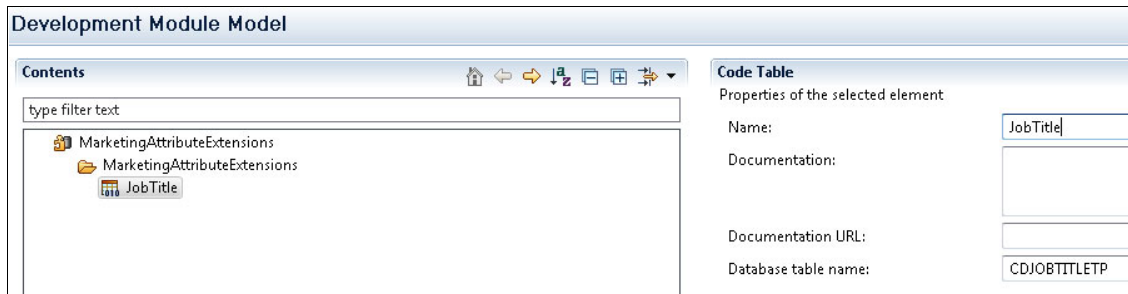
Figure 3-7   JobTitle Code Table

3. Create a new Entity Type Extension called `XPerson` by right-clicking the folder and choosing new Entity Type Extension. This holds the extra attributes to be associated with a person.

4. Specify Person as the database table for `XPerson`, as shown in Figure 3-8.
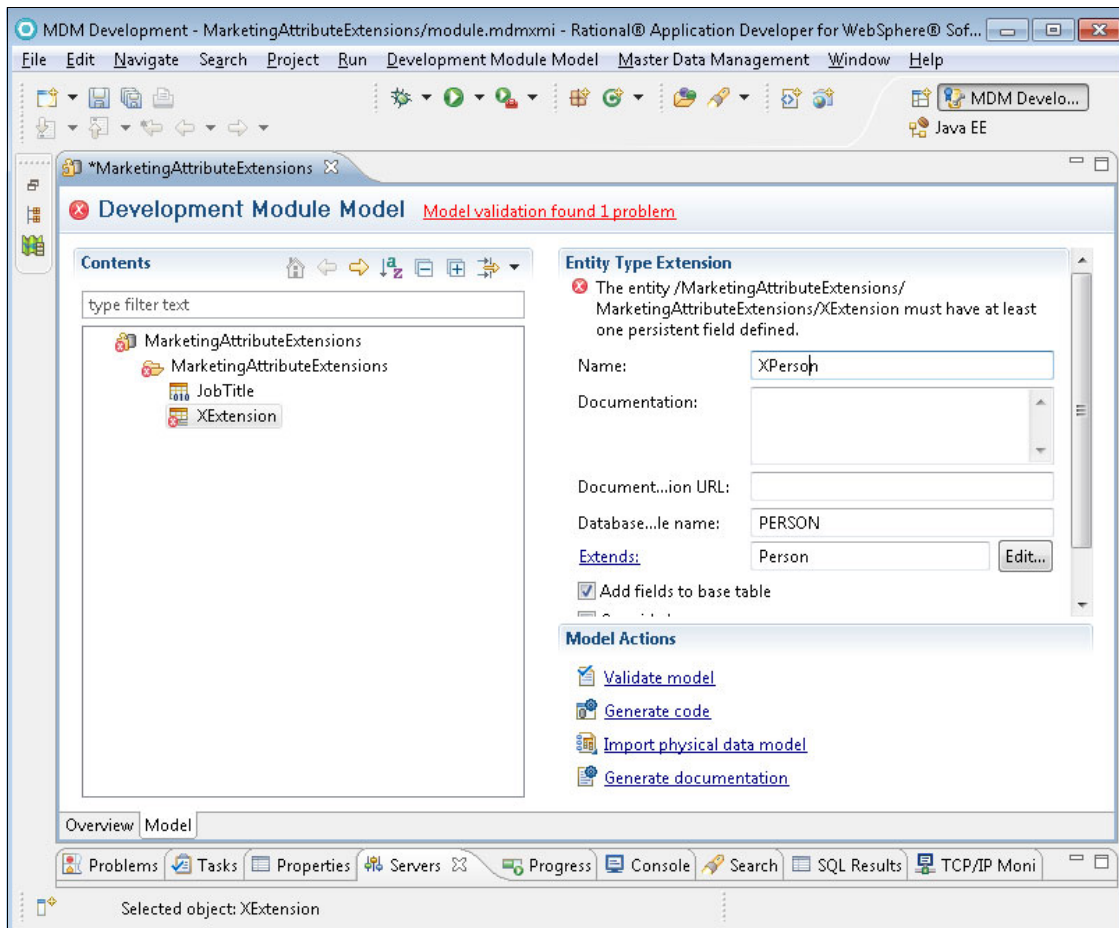
Figure 3-8   XPerson name

5. Add attributes into the Entity Type Extension by selecting **New** → **Type Code** attribute, as shown in Figure 3-9 on page 45.
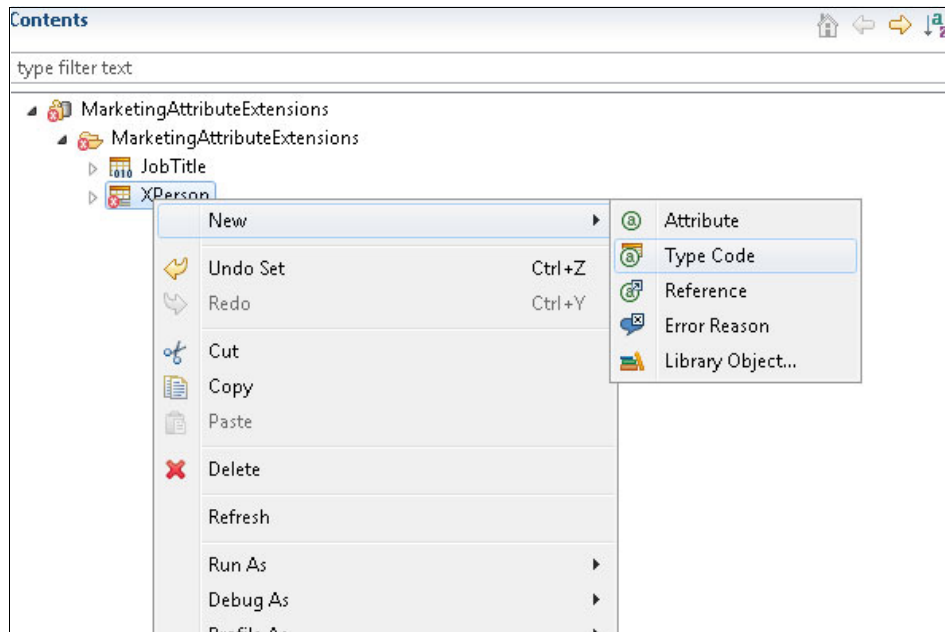
Figure 3-9   Create Type Code

6. Name the Type Code `JobTitle`. Indicate that it is nullable and refers to code table named `JobTitle`, as shown in Figure 3-10.
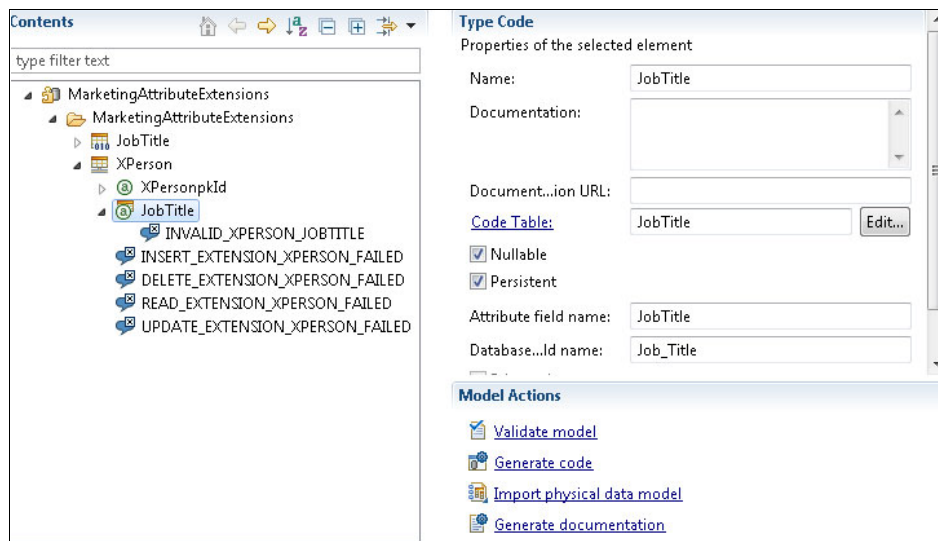
Figure 3-10   JobTitle Type Code

7. Create an attribute (this is in the menu shown in Figure 3-9 on page 45) and call it `CreditScore` with a type of integer. Then, make the attribute nullable, which means that it is not required to be specified on create or update requests. See Figure 3-11.
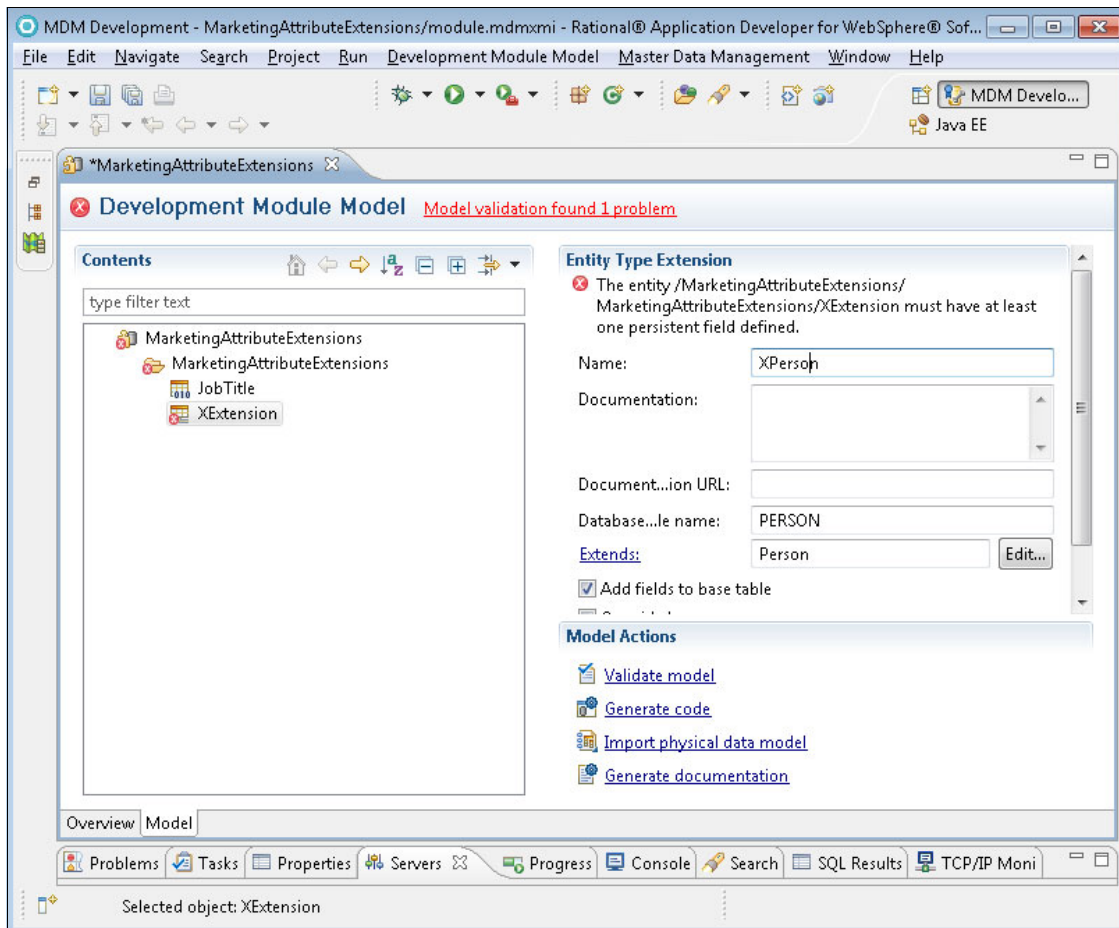
Figure 3-11   CreditScore Integer Attribute

8. Click **Generate code** to generate the code, Structured Query Language (SQL) files, and web service implementations.

Task to verify that the extension was generated

Check the XML Schema Definition (XSD) content to ensure it has picked up the customizations that have been made.

Open the `PhysicalMDM.xsd` file. You see the XSD content representing the extension, as shown in Example 3-1.

Example 3-1   XSD content representing the extension

```
<xsd:element name="XPersonBObjExt" substitutionGroup="CommonExtensionBObj"
type="XPersonBObjExtType"></xsd:element>

    <xsd:complexType name="XPersonBObjExtType">
        <xsd:annotation>
            <xsd:documentation>

            </xsd:documentation>
        </xsd:annotation>
        <xsd:complexContent>
            <xsd:extension base="CommonExtensionBObjType">
            <xsd:sequence minOccurs="1" maxOccurs="1">
                <xsd:element ref="ComponentID" minOccurs="0" maxOc-curs="1"></xsd:element>
                <xsd:element name="CreditScore" type="xsd:string" minOccurs="0"
maxOccurs="1"></xsd:element>
                <xsd:element name="JobTitleType" type="xsd:string" minOccurs="0"
maxOccurs="1"></xsd:element>
                <xsd:element name="JobTitleValue" type="xsd:string" minOccurs="0"
maxOccurs="1"></xsd:element>
                <xsd:element ref="DWLStatus" minOccurs="0" maxOc-curs="1"></xsd:element>
            </xsd:sequence>
            </xsd:extension>
        </xsd:complexContent>
    </xsd:complexType>
```

> **Note:** If you search in the MDMSharedResources for fields in the extension, it will return results only if you select the **Consider derived resources** check box on the search.

Customize the code table SQL

Several files are now generated. Ensure that the SQL used to populate the db2 code table contains the values that the marketer requires. Here, we added several example job titles into the job title table.

The SQL files that are generated for DB2 are in the MDM Development project in the `resources\sql\db2` folder. In that folder, amend the following file to replace the contents with the content shown in Example 3-2 on page 48:

`MarketingAttributeExtensions_CODETABLES_DB2.sql`

Example 3-2   Contents for the MarketingAttributeExtensions_CODETABLES_DB2.sql

```
INSERT INTO db2admin.CDJOBTITLETP ( lang_tp_cd, tp_cd, name , description,
last_update_dt, last_update_user )
   VALUES ( 100, 1, 'Account Manager/Sales Rep' , NULL, CURRENT TIMESTAMP, 'cusadmin'
);

INSERT INTO db2admin.CDJOBTITLETP ( lang_tp_cd, tp_cd, name , description,
last_update_dt, last_update_user )
   VALUES ( 100, 2, 'Associate/Individual Contributor/Specialist' , NULL, CUR-RENT
TIMESTAMP, 'cusadmin' );

INSERT INTO db2admin.CDJOBTITLETP ( lang_tp_cd, tp_cd, name , description,
last_update_dt, last_update_user )
   VALUES ( 100, 12, 'Network Administrator' , NULL, CURRENT TIMESTAMP, 'cusadmin' );

INSERT INTO db2admin.CDJOBTITLETP ( lang_tp_cd, tp_cd, name , description,
last_update_dt, last_update_user )
   VALUES ( 100, 13, 'Systems Architect' , NULL, CURRENT TIMESTAMP, 'cusadmin' );

INSERT INTO db2admin.CDJOBTITLETP ( lang_tp_cd, tp_cd, name , description,
last_update_dt, last_update_user )
   VALUES ( 100, 18, 'HR' , NULL, CURRENT TIMESTAMP, 'cusadmin' );
```

When MDM is installed, some of the code tables are already populated, often with keys that start with 1 (one). Review the existing code table content to make sure it fits with your needs. When adding new content to code tables, a good practice is to use a range of keys that differ significantly from the values of the shipped code table keys. An example of adding an entry to an existing MDM code table is as follows:

```
INSERT INTO db2admin.CDHIGHESTEDUTP ( lang_tp_cd, highest_edu_tp_cd,
name , description, last_update_dt )
   VALUES ( 100, 1001, 'High School Diploma' , NULL, CURRENT TIMESTAMP
);
```

Run the following command on a DB2 command line to populate the database. This file is in the Rational Application Developer workspace, in the MDMSharedResources\sql folder:

```
db2 -vf SETUP_DB2.sql
```

In previous versions of MDM you needed to run several SQL files in the MDM Development project the resources\sql\db2 folder.

3.6 Enhance inbound marketing with MDM

The Interact product, used for inbound marketing, is often configured to access customer information from an operational profile database. If real-time up-to-date values of the customers and prospects are needed, then Interact can call out to a web service to get this information. If you ensure that your MDM operational hub holds the up-to-date customer and prospect information, then it makes sense for MDM to expose this information as a web service. For example, if a customer has changed addresses, then calling an MDM profile service to get the up-to-date address can be a vital piece of information to capitalize on for the inbound marketing.

The MDM Workbench can be used to define exactly what pieces of prospect and customer information are exposed in the profile service. You can choose to provide some or all of the MDM operations and data in the web service. The following options can expose this profile service:

► Expose the ready-for-use MDM Party services. The advantage of this approach is that you get the person data without the need to customize MDM.

► If you want to customize how this information is exposed, use these options:

Use the MDM Adaptive Services Interface (ASI) to expose a service that has been tailored. In this way, the user decides which operations and which data will be exposed in the service. This is done using an MDM Workbench Service Tailoring Project.

If you already have some predefined services in Web Service Description Language (WSDL) format that you want to use, then MDM can be set up to expose this non-MDM WSDL. Exposing this WSDL is done in a MDM Workbench Service Mapping Project using a graphical mapper.

This book shows how to create a tailoring project that exposes only party (including person) operations. The latest address of a person can be obtained from this service.

3.6.1 Creating an MDM Service Tailoring project

In the MDM Workbench workspace, create a tailoring project then add in Party operations as shown in Figure 3-12 on page 50.
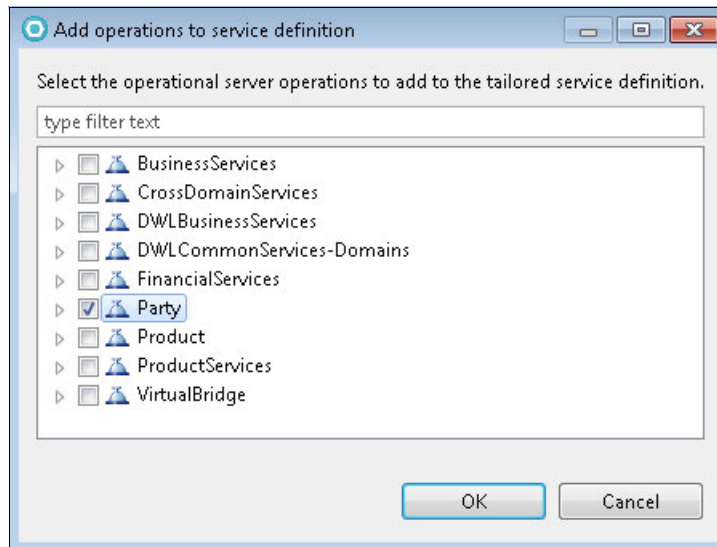
Figure 3-12   Add Party operations to service definition

To pick up the extension, you also need to add XPersonBObj to the data, as shown in Figure 3-13.
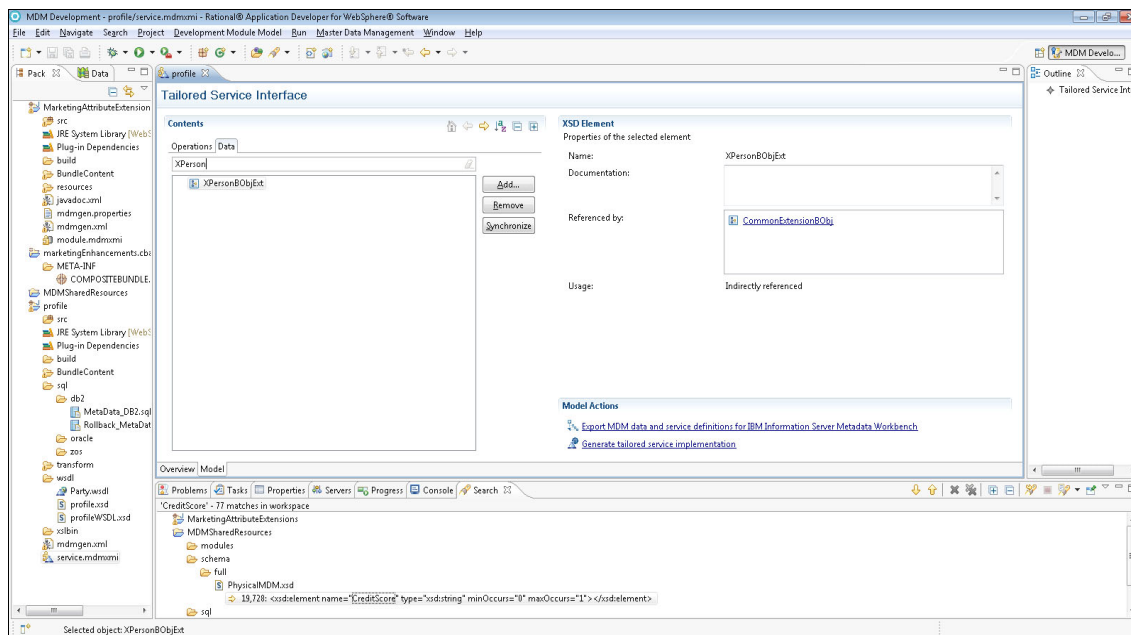
Figure 3-13   Add the Extension business object to the data

Remove any operations and data that you do not want. Then click **Generate the tailored service implementation**.

The SQL is in the `tailoring project sql/db2` folder. Run the SQL using a DB2 command shell, as in the following example:

`db2 -vf MetaData_DB2.sql`

3.7 Deploying and Testing your web services.

Now that you created the extension and profile service using the workbench, deploy these customizations to WebSphere Application Server. After they are deployed, verify that they work as expected. This section covers how to deploy the changes and how to test MDM web services.

3.7.1 Deploying a CBA

Follow the instructions in the InfoSphere MDM information center to deploy the CBA using the WebSphere Application Server administrator console:

http://pic.dhe.ibm.com/infocenter/mdm/v11r0/index.jsp?topic=%2Fcom.ibm.mdmhs.wb.tools.models.doc%2Ft_Deploy_Generated_JAX-RPC_WS.html

Then stop and start the application server. Check that the application is started by using the WebSphere Application Administration console. Start the application, if it is not already started.

3.7.2 Testing an MDM web service

One way to test a web service in Rational Application Developer is to use the web service explorer, as follows:

1. You can right-click any WSDL file in the Rational Application Developer workspace and then choose **Web Services - web service explorer**. You could run this against the tailored project WSDL or the party WSDL.

2. Choose the operation you want to test, for example search Person.

3. Then click the source link, in blue.

The web service is secure, so you need to specify security information in the SOAP header. To use the default user ID and password, specify the security information shown in Example 3-3 on page 52 in the `<soapenv.Header>` section of the SOAP header.

Example 3-3   SOAP header

```
<s:Security

xmlns:s="http://docs.oasis-open.org/wss/2004/01/oasis-200401-wss-wssecu
rity-secext-1.0.xsd"

xmlns:u="http://docs.oasis-open.org/wss/2004/01/oasis-200401-wss-wssecu
rity-utility-1.0.xsd"
    soapenv:mustUnderstand="1">
   <s:UsernameToken u:Id="unt_20">
      <s:Username>mdmadmin</s:Username>
      <s:Password

Type="http://docs.oasis-open.org/wss/2004/01/oasis-200401-wss-username-
token-profile-1.0#PasswordText">mdmadmin</s:Password>
   </s:UsernameToken>
</s:Security>
```

In the `<soapEnv.body>` section, paste in the transaction request. See the following example transaction requests:

- ► "Example searchPerson request" on page 52
- ► "Example addPerson request" on page 53
- ► "Example updatePerson request" on page 54

Example searchPerson request

Comments have been added in Example 3-4 to help explain the XML request content.

Example 3-4   searchPerson request

```
<q1:searchPerson>
  <q0:RequestControl>
    <q0:requestID>1</q0:requestID>
    <q0:DWLControl>
      <q0:requesterName>mdmadmin</q0:requesterName>
      <q0:requesterLanguage>100</q0:requesterLanguage>
    </q0:DWLControl>
  </q0:RequestControl>
  <q0:TCRMPersonSearchBObj>
    <q0:LastName>%</q0:LastName>  >   <!-- % is a wild card indicating that we are looking for
all names -'
    <q0:InquiryLevel>1</q0:InquiryLevel>  <!-inquiryLevel is a mandatory input to this request.
Specifying higher number returns more information-'
  </q0:TCRMPersonSearchBObj>
 </q1:searchPerson>
```

When you run this transaction, you see a response that contains a success return code followed by a list of all the people defined in MDM.

Example addPerson request

Example 3-5 shows an addPerson request that uses the resources defined in this chapter. We added comments to explain the XML request content.

Example 3-5   addPerson request

```
<q1:addPerson>
   <q0:RequestControl>   <!-- this is the standard MDM request control  -->
      <q0:requestID>1</q0:requestID>
      <q0:DWLControl>
         <q0:requesterName>mdmadmin</q0:requesterName>
         <q0:requesterLanguage>100</q0:requesterLanguage>   <!-this is the code table type of
language 100 is US english.-->
      </q0:DWLControl>
   </q0:RequestControl>
   <q0:TCRMPersonBObj>   <!-- this is the Person business object -->
      <q0:SolicitationIndicator>y</q0:SolicitationIndicator>   <!-- we are not opting out of
all solicitation via contact methods-->
      <q0:BirthDate>1949-08-31 00:00:00.0</q0:BirthDate>  <!-- be care of the date format-->
      <q0:GenderType>m</q0:GenderType>          <!-- the possible values are "m" or "f" and are
not in a code table -->
      <q0:HighestEducationType>1001</q0:HighestEducationType>   <!-- this refers to the code
table type we added. Specifying the type in requests is easier that specifying the value -->
      <q0:MaritalStatusType>1</q0:MaritalStatusType>   <!-- this is the code table type in
CDMARITALSTTP. The value is "Married". -->
      <q0:TCRMPartyAddressBObj>
         <q0:AddressUsageType>1</q0:AddressUsageType> <!-- this is the code table type in
CDADDRUSAGETP. The value is "Primary Residence"  -->
         <q0:TCRMAddressBObj>
            <q0:AddressLineOne>8989 N Gainey Center Dr</q0:AddressLineOne>
            <q0:City>Scottsdale</q0:City>
            <q0:ZipPostalCode>85258</q0:ZipPostalCode>
            <q0:ProvinceStateType>4</q0:ProvinceStateType> <!-- <!-- this is the code table
type in CDPROVSTATETP. Check the content makes sense for your geography-->
            <q0:CountyCode>USA</q0:CountyCode>
         </q0:TCRMAddressBObj>
      </q0:TCRMPartyAddressBObj>
      <q0:TCRMPartyContactMethodBObj>
         <q0:ContactMethodUsageType>1</q0:ContactMethodUsageType>  <!-- this is the code table
type from CDCONTMETHTP- the value is "Home Tele-phone"  -->
         <q0:SolicitationIndicator>y</q0:SolicitationIndicator>      <!-- indicate that we can
be solicited on this contact method.  -->
         <q0:TCRMContactMethodBObj>
            <q0:ReferenceNumber>(480) 411-6600 </q0:ReferenceNumber>  <!-- the phone number
-->
```

```
            <q0:ContactMethodType>1</q0:ContactMethodType>  <!-- this is the code table type
from code table CDCONTMETHCAT- the value is "Tele-phone Number" -->
          </q0:TCRMContactMethodBObj>
       </q0:TCRMPartyContactMethodBObj>
       <q0:TCRMFinancialProfileBObj>
          <q0:TCRMIncomeSourceBObj>
              <q0:IncomeSourceType>1</q0:IncomeSourceType> <!-- this is the code table type from
code table CDINCOMESRCTP - value "Annual Salary" -->
              <q0:AnnualAmount>89000</q0:AnnualAmount>
          </q0:TCRMIncomeSourceBObj>
       </q0:TCRMFinancialProfileBObj>
       <q0:TCRMPartyIdentificationBObj>
          <q0:IdentificationType>1</q0:IdentificationType> <!-- this is the code table type
from code table CDIDTP - the value is "Social Security Number" -->
          <q0:IdentificationNumber>415-22-9845</q0:IdentificationNumber>   <!--this is the
social security number.  -->
       </q0:TCRMPartyIdentificationBObj>
       <q0:TCRMPersonNameBObj>
          <q0:NameUsageType>1</q0:NameUsageType> <!-- this is the code table type  code table
CDNAMEUSAGETP - the value is "Business" -->
          <q0:StdGivenNameOne>Charles</q0:StdGivenNameOne>
          <q0:LastName>Watson</q0:LastName>
       </q0:TCRMPersonNameBObj>
    </q0:TCRMPersonBObj>
</q1:addPerson>
```

When you run this transaction, you see a response that contains a success return code. The response also contains the partyId, PersonId, and lastupdatedate, which must be specified on a subsequent updatePerson.

Example updatePerson request

Example 3-6 is an updatePerson request referring to resources defined in this chapter. We added comments to explain the XML request content.

Example 3-6   updatePerson request

```
<q1:updatePerson>
  <q0:RequestControl>
    <q0:requestID>1</q0:requestID>
    <q0:DWLControl>
      <q0:requesterName>mdmadmin</q0:requesterName>
      <q0:requesterLanguage>100</q0:requesterLanguage>
    </q0:DWLControl>
  </q0:RequestControl>
  <q0:TCRMPersonBObj>
  <!-update calls need to specify the identifier of the object being updated; in this case the
identifier is partyId. You can get the partyId from the re-sponse of a successful addPerson.
```

```
Alternatively you can search for the Person, if successful the PersonBObj is returned
containing the PartyId   -->
    <q0:PartyId>801138615542493945</q0:PartyId>
  <!-- Update calls need to specify last update date of theobject being updat-ed. For
updatePerson there is a lastupdatedate for Party and one for Person.
  You can get these from the response of the last successful add or update call. Alternatively
you can search for the Person, if successful the PersonBObj is returned containing the last
update dates-->
    <q0:PartyLastUpdateDate>2013-12-04 11:10:24.938</q0:PartyLastUpdateDate>
    <q0:PersonLastUpdateDate>2013-12-04 11:10:25.096</q0:PersonLastUpdateDate>
<TCRMExtension>  <!-- this is indicating that there is an extension to the Person   -->
  <ExtendedObject>XPersonBObjExt</ExtendedObject>   <!-- this is extension name    -->
 <XPersonBObjExt>  <!-- the extension name    -->
       <CreditScore>576</CreditScore>
      <JobTitleType>1</JobTitleType>  <!-- this is the code table type from the new code table
CDJOBTITLETP, the value is "Account Manager/Sales Rep"   -->
  </XPersonBObjExt>
  </TCRMExtension>
    <q0:TCRMPartyContactMethodBObj>
     <q0:ContactMethodUsageType>7</q0:ContactMethodUsageType>  <!-this is the code table type
from CDCONTMETHTP- the value is "Personal Email"  -->
     <q0:SolicitationIndicator>y</q0:SolicitationIndicator>  <!-- indicate that we can be
solicited on this contact method.  -->
     <q0:TCRMContactMethodBObj>
       <q0:ReferenceNumber>CWatson@yahoo.com</q0:ReferenceNumber>  <!-- email address  -->
       <q0:ContactMethodType>2</q0:ContactMethodType>    <!-this is the code table type from
code table CDCONTMETHCAT- the value is "Email Address" -->
     </q0:TCRMContactMethodBObj>
    </q0:TCRMPartyContactMethodBObj>
    <q0:TCRMPartyContactMethodBObj>
     <q0:ContactMethodUsageType>5</q0:ContactMethodUsageType> <!-this is the code table type
from CDCONTMETHTP- the value is "Cellular"  -->
     <q0:SolicitationIndicator>y</q0:SolicitationIndicator>
     <q0:TCRMContactMethodBObj>
       <q0:ReferenceNumber>(480) 167-9145 </q0:ReferenceNumber>
       <q0:ContactMethodType>1</q0:ContactMethodType>
     </q0:TCRMContactMethodBObj>
    </q0:TCRMPartyContactMethodBObj>
  </q0:TCRMPersonBObj>
</q1:updatePerson>
```

After running this transaction, you see a response that contains a success return
code. The response also contains the partyId, personId, and lastupdatedate,
which must be specified on a subsequent updatePerson.

3.8 Summary

This chapter demonstrates how to configure customer and prospect master data, accounting for marketing-relevant attributes. It describes how to install MDM and customize it using the MDM Workbench, and provides deployment and testing considerations. The marketing repository is a data mart that is created from the data warehouse for use by IBM Campaign (outbound marketing), and is populated using the data in MDM. A profile web service is provided by MDM so that IBM Interact can get real-time customer information for inbound marketing.

By following the examples in this chapter, you can create a customer master which is a solid foundation of reliable customer and prospect information, on which you can build your EMM solutions.

PureData for Analytics: The marketing data repository

Our marketing campaign scenario is based on the marketing repository deployed in an IBM PureData™ System for Analytics. In this chapter, we describe how the marketing repository can be created and provisioned in a managed and governed way. Metadata is used as the foundation to build the repository and as a result there is a consistent repository and also confidence in the data made available to the marketers. Data quality, accountability, traceability, origin, business meaning, privacy, and security are all attributes of an information supply chain that is defined and managed by using metadata, from the sources all the way to the marketing repository. InfoSphere Information Server provides the infrastructure to define and manage the information supply chain and the metadata to govern every link of it.

The remainder of this chapter describes how the marketing repository is created using a well-governed information supply chain.

For more details, see *Enhancing IBM Enterprise Market Management solutions with master data. A trusted base for EMM* at the developerWorks website:

http://www.ibm.com/developerworks/data/library/techarticle/dm-1303enhance/index.html?ca=dat-

4.1 The architecture blueprint

Figure 4-1 on page 59 depicts the information supply chain with the marketing repository as target hosted in a PureData System for Analytics. Enterprise Marketing Management (EMM) consumers, such as IBM Campaign and IBM Interact, work directly with the data in the repository and are described in Chapter 6, "Leveraging Customer Data within IBM Campaign" on page 103.

The marketing repository is sourced from an industry-specific data warehouse as described in Chapter 2, "Using IBM Industry Models for marketing" on page 13. Data moves from the data warehouse to the marketing repository with the help of the IBM InfoSphere Data Click product. There are multiple input streams to the data warehouse, most prominently the master data management system hosting the Customer Master Repository, which is introduced in 3.1, "The value of a customer master" on page 28. With the golden master records for customers and additional transactional information sources, there is a potentially long list of sources to the data warehouse. Although some of these sources can again be integrated with Data Click, others use ETL (extract, transform, load) for more complex transformations, or IBM InfoSphere Change Data Capture (CDC) for synchronized integrations with the sources.

Every aspect of this information supply chain is managed through metadata shown in the lower part of Figure 4-1 on page 59. Different types of metadata are introduced in Chapter 4, "PureData for Analytics: The marketing data repository" on page 57 and Chapter 5, "From Master Data Management to PureData for Analytics" on page 85, including physical metadata, operational metadata, lineage metadata, stewardship, policy, business and quality metadata. Although some of the metadata is used to document the information supply chain, other metadata is needed to validate the quality or consistency of data, or to drive the integration between nodes shown in Figure 4-1 on page 59.
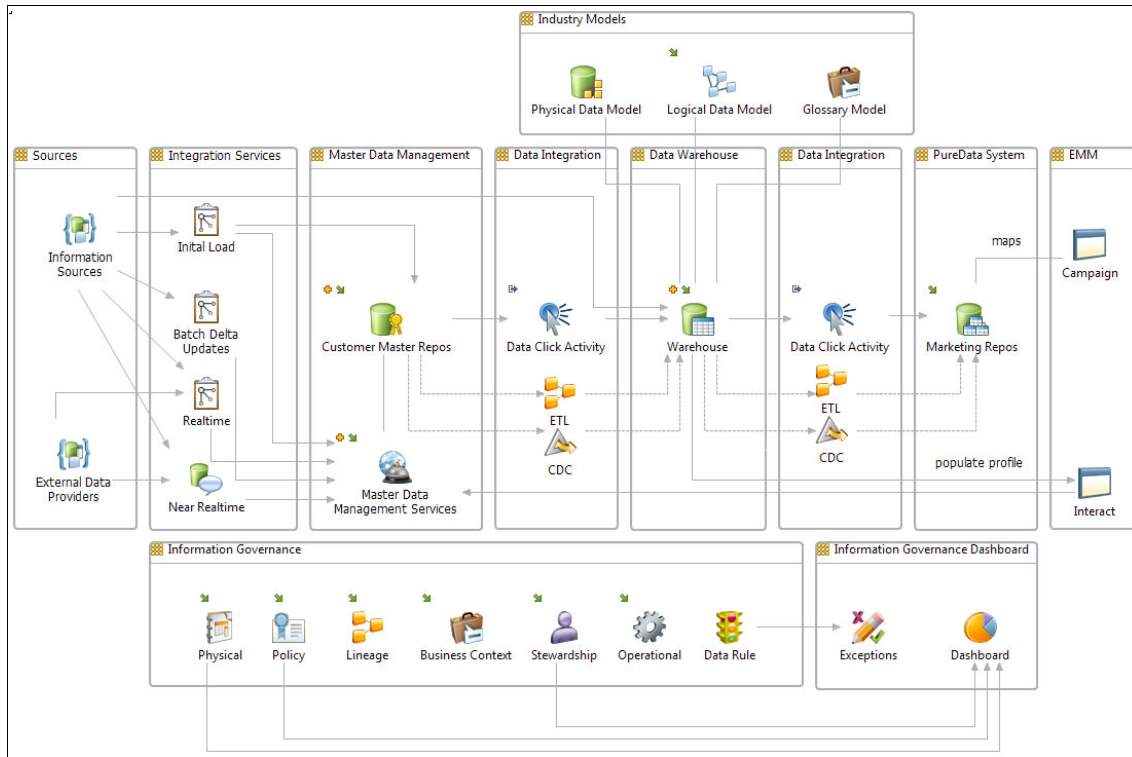
Figure 4-1   Architecture blueprint for the information supply chain

4.2 Roles and personas

Various aspects of the blueprint are described in this chapter, and also
Chapter 5, "From Master Data Management to PureData for Analytics" on
page 85 for different personas and user roles. We introduce the roles here and
continue to use them either explicitly or implicitly throughout the chapters in this
book.

► Master data developer: This is a user of the Master Data Workbench
(introduced in Chapter 3, "Configuring MDM with marketing-relevant
attributes" on page 27.) The developer builds custom models, service tailoring
projects, and facilitates the export of such models.

► Database administrator: This role refers to one or multiple individuals who
configure the master data hub used for the customer master repository, the
data warehouse, and the data marts in the PureData System for Analytics.

- Metadata administrator: This is the role of a person working with the InfoSphere Metadata Workbench to import metadata or modify existing metadata. The administrator is assumed to have privileges for all of the metadata captured in the Information Governance area of the blueprint.

- Data click administrator: This role works with InfoSphere Blueprint Director to configure Data Click using metadata. This role can be merged with the role of the metadata administrator for simplicity.

- Business user: This is any user working in any line of business. That user has an understanding of business concepts and processes and will use a typical business terminology. The technical nature of the data that this user works with is not known, but the user is able to understand and navigate the information supply chain with tools that are simple to use.

- Self-service user: This is a business user with the special requirement to access physical data in either the Customer Master Repository or the data warehouse. This might be a business analyst, a data analyst, or a marketing person working on the marketing campaign. A self-service user does not possess technical skills such as ETL or SQL and does not have any database credentials.

- Data steward: This is an individual responsible for the quality and integrity of some aspects of the information supply chain. The scope varies based on the maturity of an organization. Some Data Stewards are responsible for all of the data in the data warehouse, including the loading of ETL processes; others are in charge of only one particular field in one table. Data Stewards are responsible for any exception in their managed data or data flows.

- Governance officer: This is the individual responsible for governance aspects of the organization. In the context of this book, we consider only data governance as an aspect of IT governance and ignore corporate governance altogether. Even within the reduced scope of data governance, businesses often create hierarchies of governance officers reporting all the way up to a C-level officer.

4.3 Data integration

Data integration has been studied for decades and several approaches have been developed, including extract, transform, and load (ETL), replication, data virtualization, federation, or linked open data. Our marketing campaign leverages a traditional ETL style integration and data is physically copied from sources on the left to targets on the right.

We do, however, expand the definition of data integration to include Data Click feature.

Data Click is an implementation of what is sometimes referred to as "shop-for-data" in the literature. Instead of unmanaged individual data copies, consider the concept of a "shopping cart" that users can fill and then execute the "checkout" operation. The process is managed and governed by policies, similar to any supermarket purchase. Policies are in place to control what data can be provisioned to the target system, how much can be provisioned, or who is allowed to provision it. Every transaction is monitored and persisted in a transaction history to make it fully traceable. InfoSphere Information Server provides the metadata repository used to implement the "shop-for-data" integration and the technology stack needed to execute it.

The following two main data integration links are described in this chapter:

▶ Integration of the Customer Master Repository with the data warehouse
▶ Integration of the data warehouse with the marketing repository

Both integrations can be implemented using one of the following two approaches:

▶ Data Click to "shop for data" in the source and load it into the target without any transformation
▶ IBM InfoSphere DataStage® ETL jobs to extract data from the source, perform transformations, and load data into the target

IBM InfoSphere DataStage ETL is the more flexible and generic approach but might be less appealing to the self-service user because it requires ETL development skills, whereas Data Click can offer a simple two-click user experience.

Data Click can be more simply governed than IBM InfoSphere DataStage and an administrator can limit the volume of data being moved and the load a user puts on the system. Both technologies are metadata-driven, where metadata is used for the configuration during design time and produced upon execution at run time. A replication strategy is available with Data Click and IBM InfoSphere DataStage, which allows you to continuously refresh the target with changes from the source.

The decision of what technique to use depends on the requirements and the circumstances available at the given link in the information supply chain.

If, for example, the Customer Master Repository is to be integrated with a new data warehouse, you can use Data Click for a simple and governed integration without any transformation requirements in that case.

If, however, the data warehouse exists and customer records have to fit a given schema in the data warehouse or be joined with other transactional data, you use IBM InfoSphere DataStage ETL jobs to facilitate the transformations during load.

The remainder of the chapter introduces both options for integration between the Customer Master Repository and the warehouse. The same concepts apply for the integration between the data warehouse and the marketing repository.

4.4 Metadata

Both integrations are done by the use of metadata that is made available through a metadata layer. Metadata is used to drive the execution of a Data Click activity or an IBM InfoSphere DataStage job; alternatively, more metadata is produced during the execution itself. This ultimately becomes a connected cycle where metadata production and consumption go together.

Starting with a clean metadata repository, however, you must do an initial discovery of the customer master data repository and prepare for the first Data Click activity or IBM InfoSphere DataStage job execution. This section focuses on the minimal metadata configuration required for Data Click and IBM InfoSphere DataStage. We omit policy, stewardship, and business metadata for now (more detail about those subjects is in Chapter 5, "From Master Data Management to PureData for Analytics" on page 85). Although these types of metadata assets are all vital from a governance perspective, we reduce the scope of this chapter to the physical, operational and lineage metadata needed to configure Data Click and IBM InfoSphere DataStage.

4.4.1 Physical metadata

The only bit of physical metadata required for the configuration of Data Click is the data schema of the customer master data repository. This can be obtained by using one of the following two options:

► Option 1: Through discovery (or reverse engineering) of the master data database. This option is available in the Information Server console shown in Figure 4-2 on page 63. A metadata administrator performs a selective discovery and import of the relevant master data tables. Physical metadata imported using this method is immediately available to Data Click.
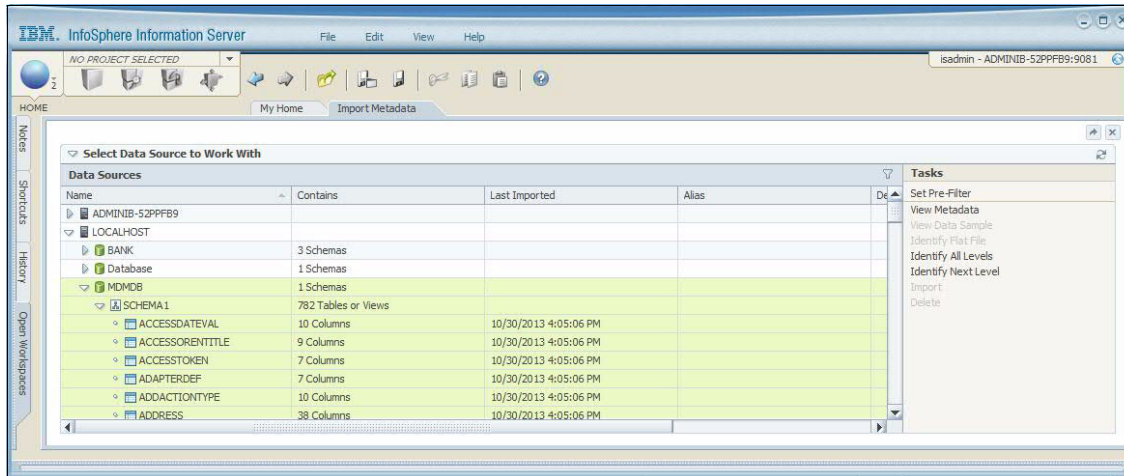
Figure 4-2   Discovery of master data schema

► Option 2: Through import of a physical database model design for the master data database. This option is available with the InfoSphere Metadata Asset Manager as part of Information Server. A metadata administrator creates an import area to use the IBM InfoSphere Data Architect MetaBroker® to import InfoSphere Data Architect models.

Using the connector to import the database model file (.dbm) for the customer master repository physical database model, the user can select **Create the database schema from physical model**. Then, the database schema will be derived from the database model. The schema will not be immediately available to Data Click. First, it must be analyzed, previewed, and published to the shared metadata repository. Unlike option 1, option 2 allows for consolidation of physical metadata during reimporting. Figure 4-3 on page 64 shows the result of a successful import using InfoSphere Metadata Asset Manager.
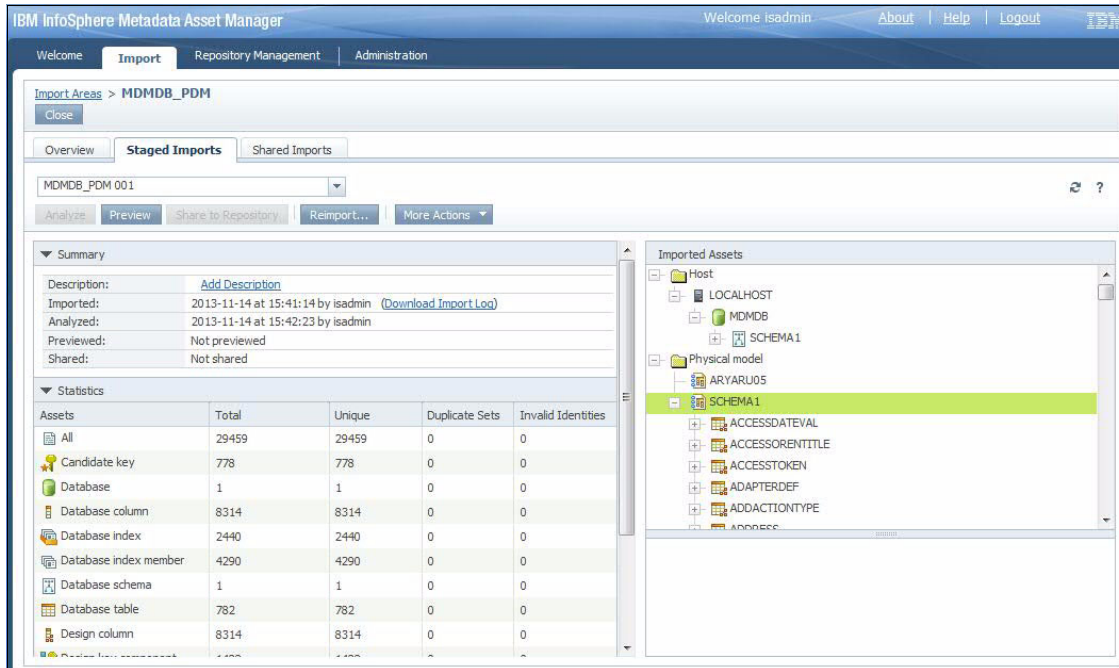
Figure 4-3   Import of customer master repository: physical schema

IBM InfoSphere DataStage ETL jobs also require an import of the data warehouse physical model. It follows the same basic flow as the physical database model import for the Customer Master Repository.

4.4.2 Lineage metadata

At this point we have made the entire Customer Master Repository data schema available in metadata. With the correct connection information we would have enough information to copy the entire Customer Master Repository database to the data warehouse. In reality, however, we need a more tailored approach and only want to copy certain tables containing the relevant attributes in the Customer Master Repository to the data warehouse. This section describes how to reduce the scope of the data integration and operate only on a filtered data set. This is achieved by sharing lineage metadata from the source Customer Master Repository with InfoSphere Metadata Workbench.

Export lineage metadata

First, we again use the MDM Workbench (used in Chapter 3, "Configuring MDM with marketing-relevant attributes" on page 27). Reopen the service tailoring project and in the Model tab find the Model Action Export MDM data and service

definitions for IBM Information Server Metadata Workbench. If there is no service tailoring project to work with, the export wizard can be selected from the Export menu.

Figure 4-4 shows the first page of the export wizard. In the fields in the Database connection section, provide the database connection to the Customer Master Repository database and the default schema used for the repository tables.

Figure 4-4   MDM Workbench export wizard page one

If a service tailoring project is available or the wizard has been launched from within such a project, it can be used directly to filter the export and include only the service and object definitions from the project.

The export consists of the following two files; both files are produced into the same folder:

► A single file containing all of the following items:

    – The repository services
    – The repository service methods
    – The repository data objects

► A mapping file that maps the following items:

    – Repository services to service methods
    – Repository service methods to input and output parameters
    – Input and output parameters to repository data objects
    – Repository data objects to tables
    – Object attributes to columns

A small second wizard page (Figure 4-5) requests the following string substitution parameters:

► Application name: The name for a container used to populate all repository services and objects into. It serves as a namespace in metadata server.

► Host name: The name of the server that hosts the Customer Master Repository database. It is important to match this host name with the name of the discovered database in section 4.4.1, "Physical metadata" on page 62.

► Database name: The name of the repository database itself. It is again important to match this to the database name discovered in section 4.4.1, "Physical metadata" on page 62.

> **Note:** Both host name and database name become part of the identity for every table used in a mapping (for example, address maps to local host, MDM, ADDRESS, as shown in Figure 4-5).
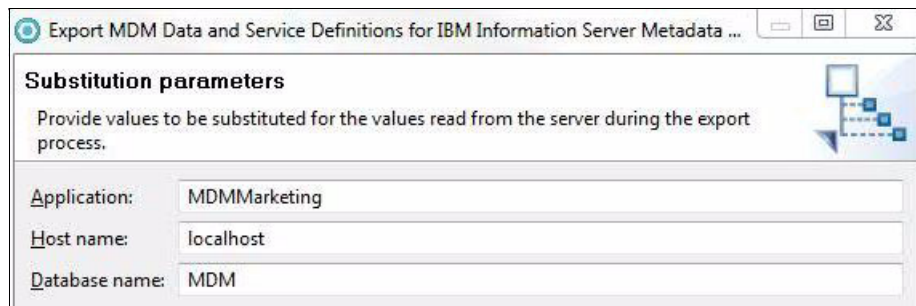

Figure 4-5   MDM Workbench export wizard page two

Both the single file and the mapping file can be inspected and also modified manually upon export. They are now ready for import into InfoSphere Metadata Workbench.

Import lineage metadata

The two exported files must be imported in the correct order:

1. First, the object definitions are loaded using the Import Extended Data Sources wizard under the Advanced tab.

   Select the first file exported with the MDM Workbench wizard shown in Figure 4-5 on page 66 (for example `MDMExtendedDataSource_Marketing.csv`). The import will result in a new application type (for example MDMMarketing), which can now permanently be browsed and discovered with InfoSphere Metadata Workbench.

2. To import the second file, use the Import Extended Data Sources wizard again in the Advanced tab of Metadata Workbench. Load the `MDMExtensionMapping_Marketing.csv` file from the export location (Figure 4-4 on page 65).

Monitor the warnings during import and make sure they are acceptable. The main risk here is that of incorrect host name or database name values in "Export lineage metadata" on page 64. That will result in a large number of warnings but does not fail the import itself. Without the correct table mappings, you will not be able to use the lineage metadata effectively. Ensure that only minor warnings exist during this part of the import.

Browse lineage metadata

There are several ways to use the imported lineage metadata. It can be used for documentation purposes, for governance and provenance, or for data integration such as ours. All of these solutions are based on the same connected lineage graph. The graph is computed by Metadata Workbench over the metadata we just imported. One way to understand the graph is through visualization.

To visualize the entire graph, find the MDMMarketing application with the asset browser. At a first level, it contains Objects and Services, under Object Type as shown in Figure 4-6.
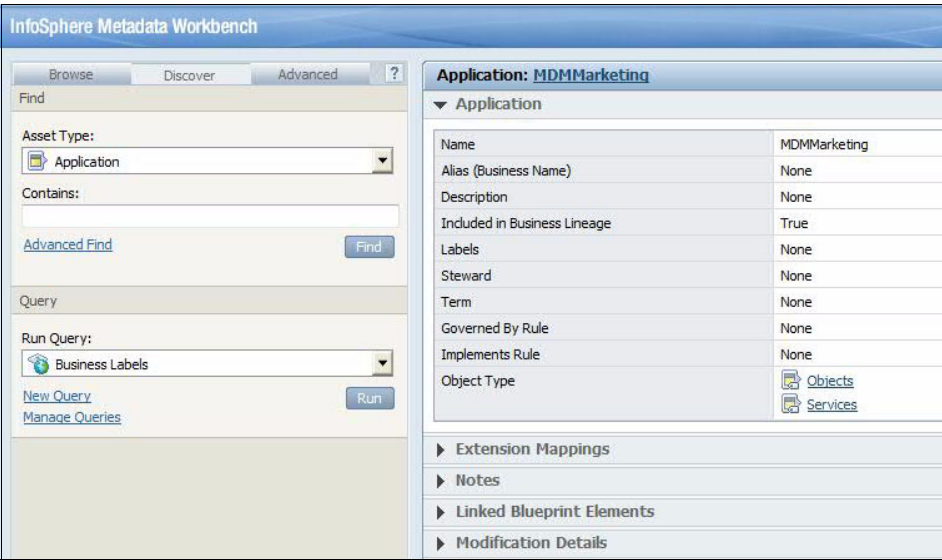

Figure 4-6   MDM Marketing application in Metadata Workbench

Open any object or service in the list and select Business Lineage. The complete graph will be rendered starting with the selected object or service, as shown in Figure 4-7 on page 69. It visualizes all descendants according to the mappings exported in "Export lineage metadata" on page 64.
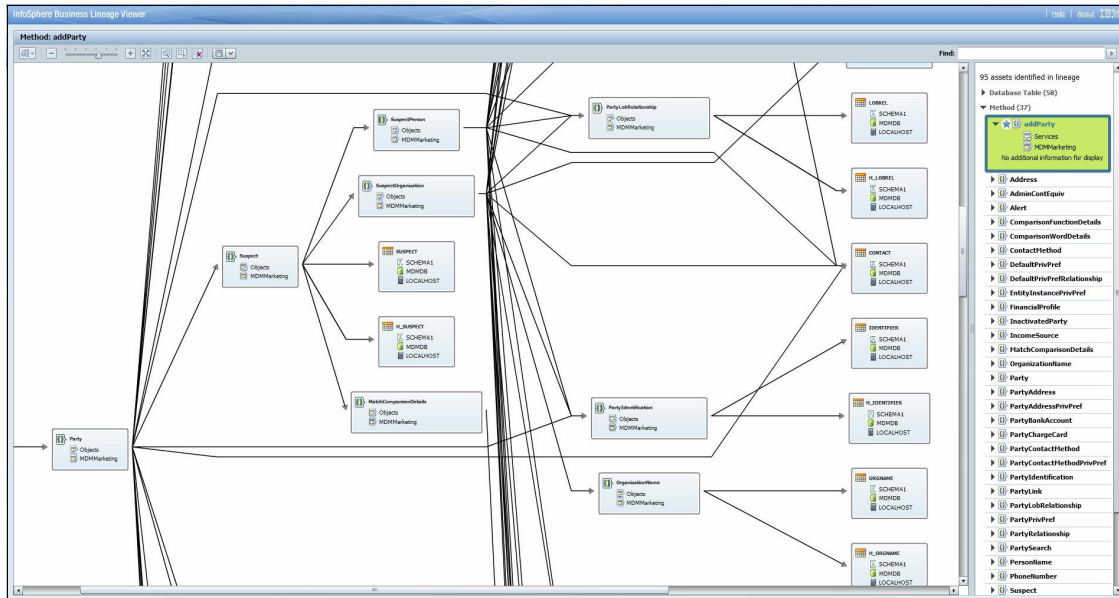
Figure 4-7   MDM lineage metadata graph

Use lineage metadata

Finally, you can use the lineage graph to determine the set of tables and columns needed within the data warehouse from our Customer Master Repository. A list of objects and attributes needed for the marketing campaign is in 3.5, "Adding marketing attributes to MDM" on page 37 already provided. You can now complete that list by inspecting the lineage metadata.

Instead of working with the complete visualized graph, you can query the metadata server using the correct asset type in the following list:

► For service names query asset type: Method (for example, addAddress)

► For object names query asset type: Method (for example, Address)

► For service input parameters query asset type: Input Parameter (for example, addAddress(Address))

► For service output parameters query asset type: Output Parameter (for example, getAddress() = Address)

► For table names query asset type: Database Table (for example, ADDRESS)

► For column names query asset type: Database Column (for example, CITY_NAME)

Figure 4-8 shows a PersonName object with all of its attributes. Select an individual attribute from the list and click **Impact Analysis**.
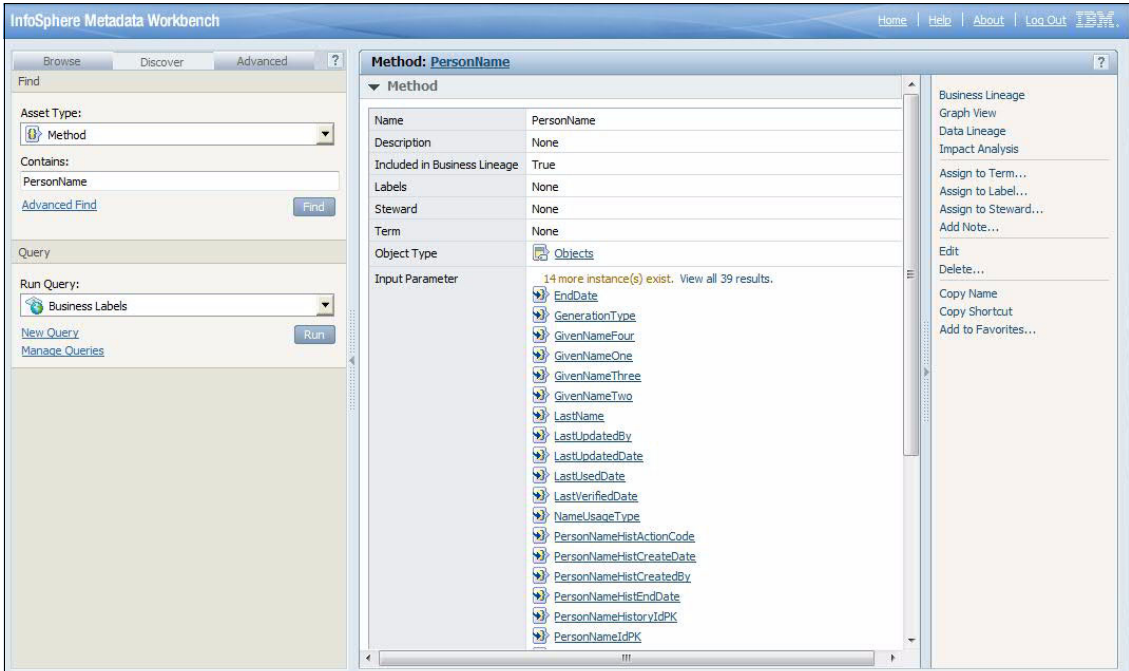

Figure 4-8   PersonName metadata

This will result in a small graph showing only the columns directly mapped by the object attribute, as shown in Figure 4-9.
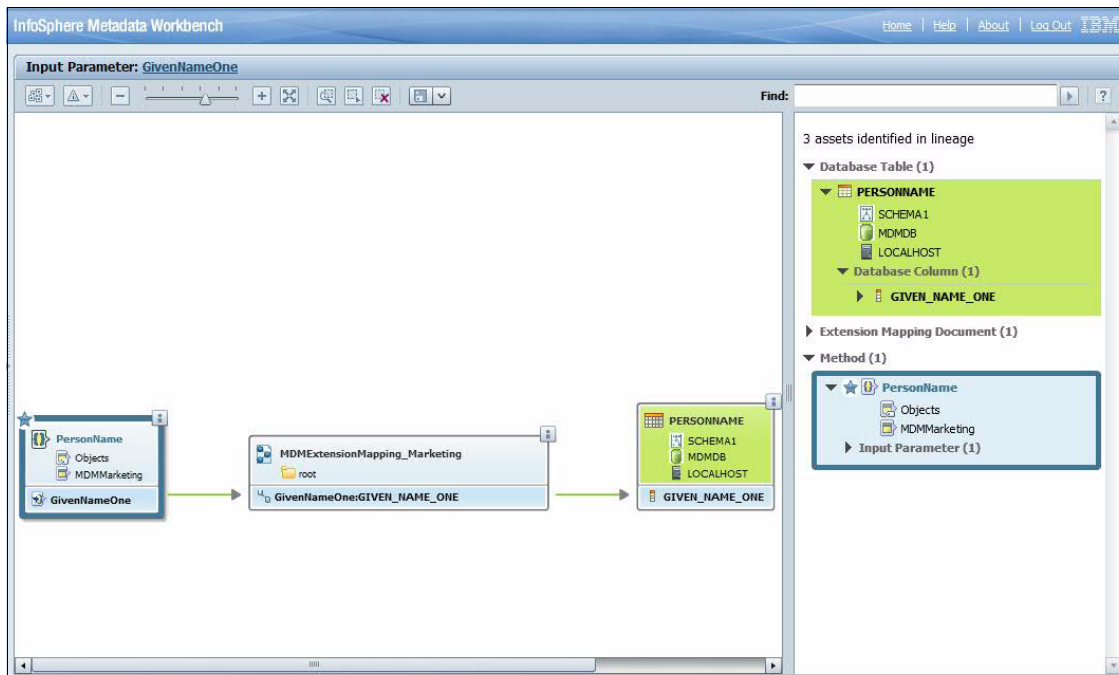

Figure 4-9   Impact analysis for PersonName attribute GivenNameOne

For a single attribute, GivenNameOne, you see how it maps to a `GIVEN_NAME_ONE` column in `PERSONNAME` table in `LOCALHOST.MDMDB.SCHEMA1` schema.

This is the column to offload to the data warehouse from the Customer Master Repository to get the qualified first name of a person. This lineage-based approach allows you to complete the data set definition for the Data Click offload operation or the IBM InfoSphere DataStage ETL job.

Table 4-1 is the result of queries against our lineage metadata for the project properties from Chapter 3, "Configuring MDM with marketing-relevant attributes" on page 27.

Table 4-1   Extracted master customer repository fields

Attribute name	MDM db table	MDM table column
Source system code	CONTEQUIV	ADMIN_SYS_TP_CD
Source system record ID	CDIDTP	ID_TP_CD

Attribute name	MDM db table	MDM table column
First name	PERSONNAME	GIVEN_NAME_ONE
Middle name	PERSONNAME	GIVEN_NAME_TWO
Last name	PERSONNAME	LAST_NAME
Address Line 1	ADDRESS	ADDR_LINE_ONE
Address Line 2	ADDRESS	ADDR_LINE_TWO
City	ADDRESS	CITY_NAME
State/Province	ADDRESS	PROV_STATE_TP_CD
Post code	ADDRESS	POSTAL_CODE
Country	ADDRESS	COUNTRY_TYPE_CD
Social Insurance ID number	CONTEQUIV	ADMIN_CLIENT_ID
Birth date	PERSON	BIRTH_DT
Income	INCOMESOURCE	ANNUAL_AMT
Gender	PERSON	GENDER_TP_CODE
Language code	PERSON	PREF_LANG_TP_CD
Marital Status	PERSON	MARITAL_ST_TP_CD
Education	PERSON	HIGHEST_EDU_TP_CD
Home phone	CONTACTMETHOD	REF_NUM
Email address	CONTACTMETHOD	REF_NUM
Mobile phone	CONTACTMETHOD	REF_NUM
Mail Opt In	LOCATIONGROUP	SOLICIT_IND
Phone Opt In	LOCATIONGROUP	SOLICIT_IND
Email Opt In	LOCATIONGROUP	SOLICIT_IND
Opt Out of All	CONTACT	SOLICIT_IND
Preferred channel	LOCATIONGROUP	PREFERRED_IND
Preferred channel ID	CDMETHODSTATUSTP	METHOD_ST_TP_CD

4.5 Data Click

We first describe our integration using Data Click and follow with a description of IBM InfoSphere DataStage ETL.

4.5.1 Preparing the Data Click execution

Data Click requires physical connections to the source and target systems used in the "shop-for-data" implementation. According to the blueprint (Figure 4-1 on page 59), you require operational database connections for the Customer Master Repository and the data warehouse in the first Data Click activity, and connections to the data warehouse and the marketing repository in the second Data Click activity.

Database connections are provided to Data Click in the blueprint that contains the Data Click activity. The blueprint has been built with InfoSphere Blueprint Director, a foundation tool in the suite of InfoSphere products.

For the blueprint shown in Figure 4-1 on page 59, you can open the Asset Browser view and list the physical data schema assets imported in 4.4.1, "Physical metadata" on page 62.

Figure 4-10 on page 74 shows the Asset Browser with the previously discovered Customer Master Repository schema. A simple drag-and-drop action onto the database icon in the blueprint associates the metadata with the icon. Repeat this action for any other source or target used by Data Click.
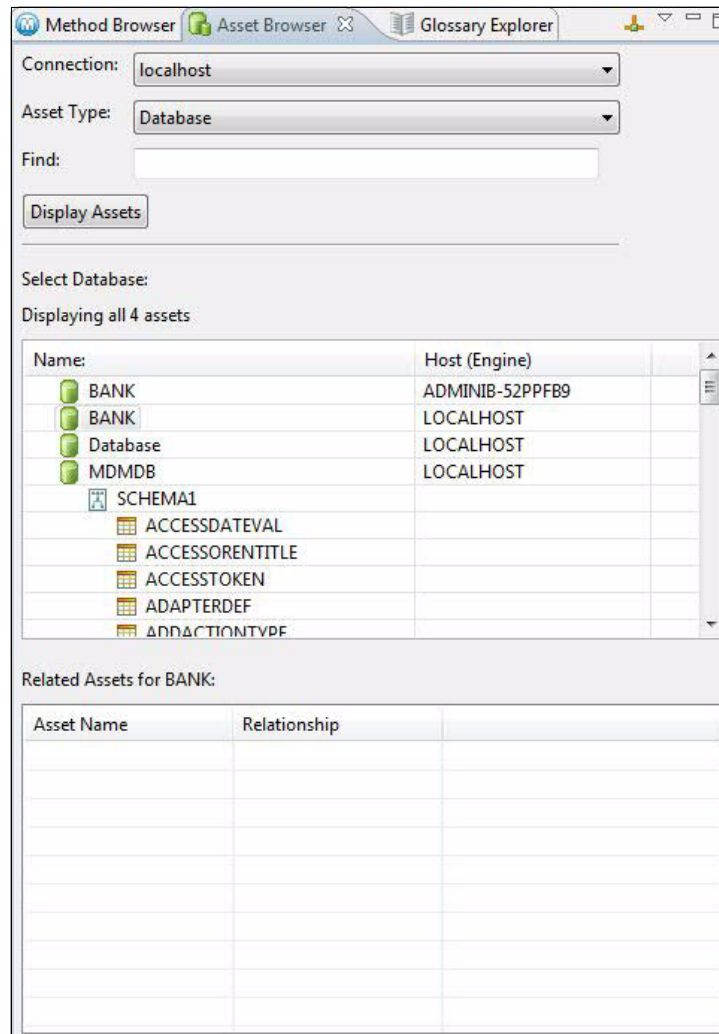
Figure 4-10   Asset Browser with physical metadata

Data Click requires a few additional pieces of information. To configure them use the **Configure InfoSphere Data Click** action in the context menu for every Data Click activity, such as the following steps:

1. Provide the Information Server connection that will be used for execution.

2. Select the schema containing the tables to be copied from the source side to the schema to be pasted in on the target side.

3. Add the location of the DataStage engine used to perform the Data Click activity.

4. Specify several configuration options to limit the control the Data Click operation.

The operational configuration of Data Click is now complete. The blueprint gets published now, to make it actionable for the self-service users. Save the blueprint and select **File** → **Publish Blueprint**. The Publish Blueprint window opens (Figure 4-11).
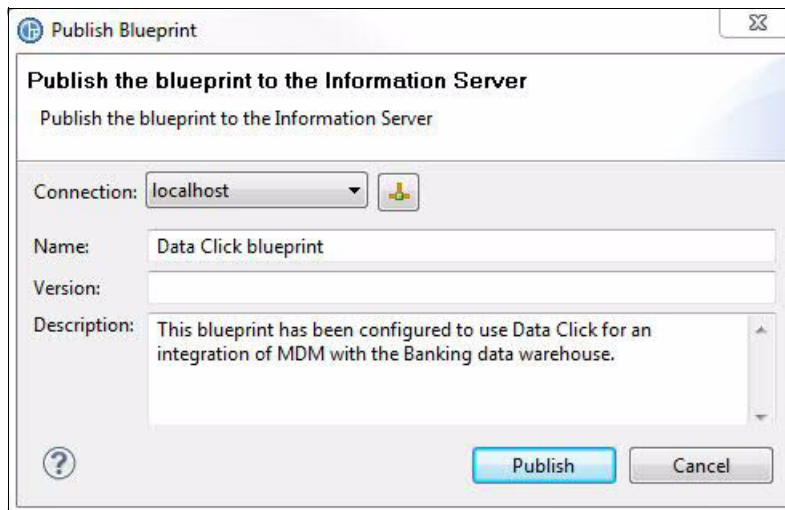

Figure 4-11   Publish Data Click blueprint

Upon publishing, you can immediately open the blueprint in the Blueprint Viewer web application available with InfoSphere Information Server.

4.5.2 Executing a Data Click activity

For self-service users to get started with the blueprint, you can either provide them with the URL or find the blueprint in Metadata Workbench. In the Metadata Workbench homepage, select **Blueprint** for Asset Type and find the blueprint with the published name.

When the business user obtains the link to the published blueprint, the blueprint can be seen in a web-based blueprint viewer similar to Figure 4-12. A single link on the right side becomes active when a Data Click Activity element is selected in the blueprint. By using this link, the Data Click configuration user interface (UI) can be launched.
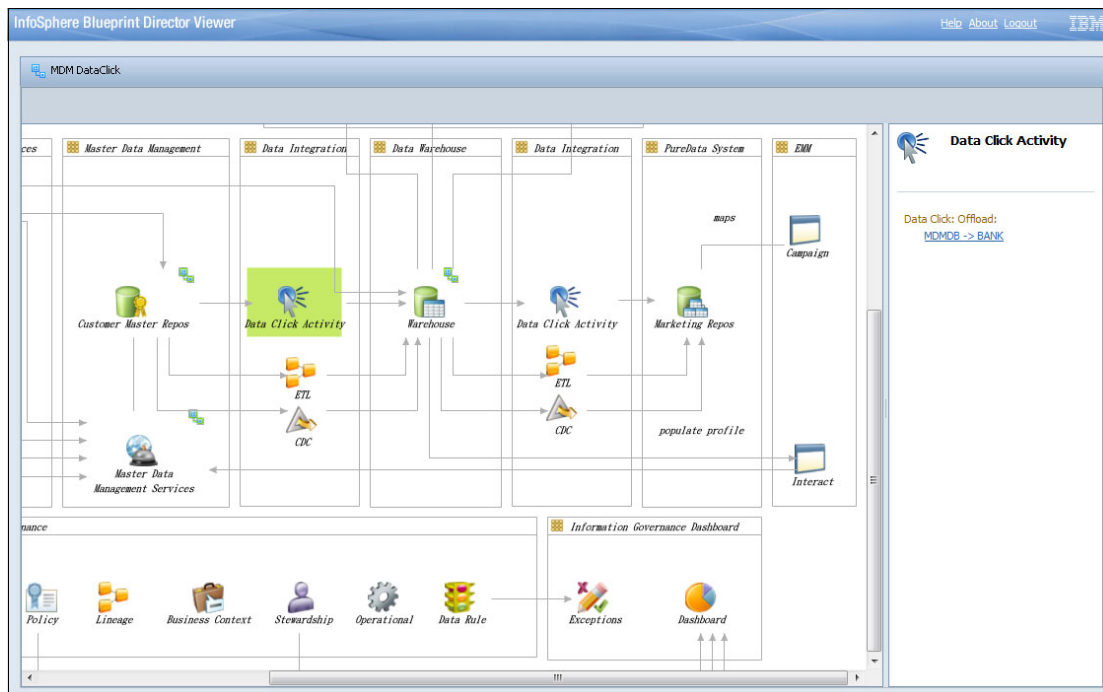

Figure 4-12   Data Click enabled web blueprint

Figure 4-13 on page 77 shows the UI as launched from the blueprint. The UI is implemented as a simple wizard with four pages to perform the following tasks:

► Select the target.
► Select the source.
► Configure the options.
► Review a summary and execute.

The Select Target page (the first wizard page) provides information to the self-service user about the configured target. It shows what tables already exist and makes the user aware if existing tables will potentially be replaced. If the tables do not exist, as in our scenario, they will be created and populated by the Data Click activity.

On the Select Source page, the user selects the columns and tables relevant for the target. In our scenario, that list is in "Use lineage metadata" on page 69, has been obtained by using the lineage metadata.
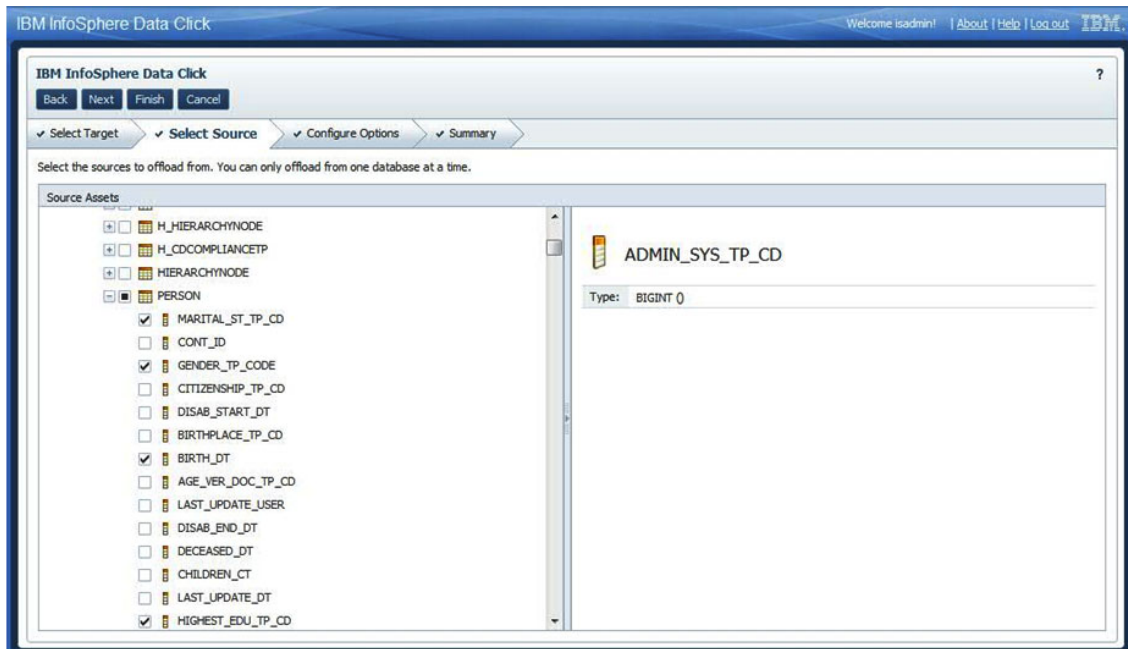

Figure 4-13   Data Click source selection

The Configure Options page in the wizard provides additional configuration options to the self-service user. We decided to then do the following tasks:

► Do an initial load of the data warehouse with data from the customer master repository only.

► Prefix the table names with MDM.

► Replace tables that potentially exist already (although we know based on the target selection that no conflict exists).

The Summary page (Figure 4-14 on page 78) in the wizard summarizes the selections made. By clicking **Offload**, the Data Click activity is triggered.

Figure 4-14   Data Click Summary configuration options

That concludes the Data Click activity for the self-service user. The Data Click UI, the blueprint, or even the web browser can now be closed. Data Click executes entirely asynchronously and in the background.

4.6 DataStage

In some situations, Data Click might not be the best option to load the data warehouse or marketing repository; several disadvantages are listed in 4.3, "Data integration" on page 60. Although the convenience of a self-service approach to data integration is beneficial, you might require more complex transformations in the integrations than those that can be expressed in simple SQL views.

If that is the case, we suggest an ETL style data integration where an ETL developer can harden the transformation logic into a DataStage job. This section describes how a job can be designed for the integration from the Customer Master Repository to the data warehouse. The exact same ideas can be applied between the data warehouse and the marketing repository. You can also mix and match and use Data Click for one of the integrations and DataStage for the other based on requirements.

4.6.1 Preparing the DataStage job execution

Figure 4-15 shows the simplest job design with a single source connector stage on the left, a target connector stage on the right, and a transformer stage in the middle. The source and target connectors must be configured with the operational connections to the source and target databases. This is done directly with the properties of the stage, and the connection information can be shared with other stages or other jobs within the project.
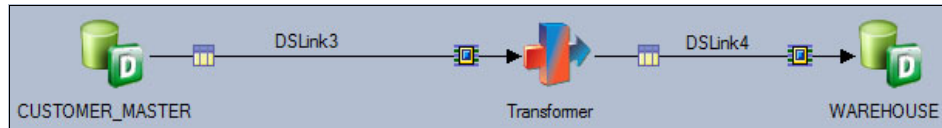
Figure 4-15   Data stage job with single transformation

Unlike the Data Click use case, you must have table definitions available for both the source and the target. The Transformer stage is used to build the mappings between the source and the target and cannot be designed without a full column listing on both sides.

You already have the Customer Master Repository physical database schema available (work accomplished in 4.4.1, "Physical metadata" on page 62). You also have the data warehouse physical schema available based on the metadata imported from the Industry Model (in 2.5, "Use the BDW to model your data requirements" on page 22).

Both physical database schema imports created what are considered shared table objects in the metadata server. For DataStage to work with the table objects, however, you must first create concrete table definitions for the required tables. Figure 4-16 on page 80 shows the wizard.

Figure 4-16   Create table definition wizard

To decide what tables are needed on the source and the target side, revisit the tables made available in the Customer Master Repository (in Table 4-1 on page 71) and the Banking Data Warehouse (BDW) Entities identified in 2.5, "Use the BDW to model your data requirements" on page 22.
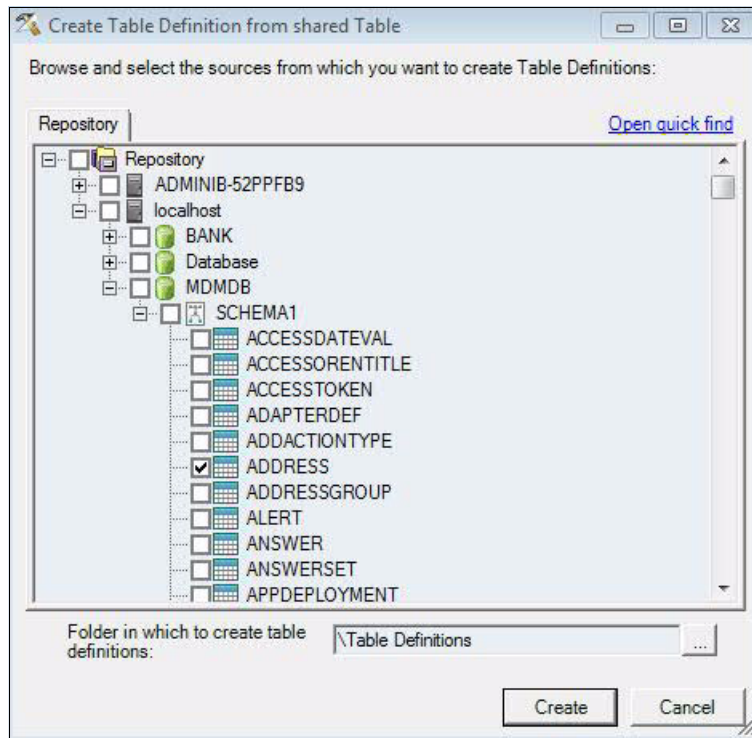
With the help of all the available lineage metadata, you can derive the table shown in Figure 4-17.

Attribute name	MDM db table	MDM table column	BDW table name	BDW table column	BDW Entity Name	BDW Attribute Name
Source system code	CONTEQUIV	ADMIN_SYS_TP_CD	IP	SRC_STM_ID	Involved Party	Source System Id
Source system record ID	CDIDTP	ID_TP_CD	IP	UNQ_ID_IN_SRC_STM	Involved Party	Unique Id In Source System
First name	PERSONNAME	GIVEN_NAME_ONE	IDV	GVN_NM	Individual	Given Name
Middle name	PERSONNAME	GIVEN_NAME_TWO	IDV	MDL_NM	Individual	Middle Name
Last name	PERSONNAME	LAST_NAME	IDV	SURNAME1	Individual	Surname1
Address Line 1	ADDRESS	ADDR_LINE_ONE	PST_ADR	ADR_NM	Postal Address	Address Name
Address Line 2	ADDRESS	ADDR_LINE_TWO	PST_ADR	STR_NBR, STR_NM	Postal Address	Street Number, Street Name
City	ADDRESS	CITY_NAME	PST_ADR	CITY_ID	Postal Address	City Id
State/Province	ADDRESS	PROV_STATE_TP_CD	PST_ADR	STE_ID	Postal Address	State Id
Post code	ADDRESS	POSTAL_CODE	PST_ADR	PSTCD_AREA_ID	Postal Address	Postal Barcode
Country	ADDRESS	COUNTRY_TYPE_CD	PST_ADR	CTY_ID	Postal Address	Country Id
Social Insurance ID number	CONTEQUIV	ADMIN_CLIENT_ID	IDV	SOC_SCR_NBR	Individual	Social Security Number
Birth date	PERSON	BIRTH_DT	IDV	BRTH_DT	Individual	Birth Date
Income	INCOMESOURCE	ANNUAL_AMT	IDV	<IDF_INCOME_AMT>	Individual	Individual Income Amount
Gender	PERSON	GENDER_TP_CODE	IDV	GND_ID	Individual	Gender Id
Language code	PERSON	PREF_LANG_TP_CD	IDV	LNG_ID	Individual	Language Id
Marital Status	PERSON	MARITAL_ST_TP_CD	IDV	IDV_MAR_ST_TP_ID	Individual	Individual Marital Status Type Id
Education	PERSON	HIGHEST_EDU_TP_CD	IDV	CTF_TP_ID	Individual	Certification Type Id
Home phone	CONTACTMETHOD	REF_NUM	TEL_ADR	TEL_ADR_ID	Telephonic Address	Telephonic
Email address	CONTACTMETHOD	REF_NUM	EMAIL_ADR	EMAIL_ADR_ID	Email Address	Email
Mobile phone	CONTACTMETHOD	REF_NUM	TEL_ADR	TEL_ADR_ID	Telephonic Address	Telephonic
Mail Opt In	LOCATIONGROUP	SOLICIT_IND	<CP>	<MAIL_OPT_IN>	Contact Preferences	Mail Opt In
Phone Opt In	LOCATIONGROUP	SOLICIT_IND	<CP>	<PHONE_OPT_IN>	Contact Preferences	Phone Opt In
Email Opt In	LOCATIONGROUP	SOLICIT_IND	<CP>	<EMAIL_OPT_IN>	Contact Preferences	Email Opt In
Opt Out of All	CONTACT	SOLICIT_IND	<CP>	<ALL_OPT_OUT>	Contact Preferences	All Marketing Opt Out
Preferred channel	LOCATIONGROUP	PREFERRED_IND	<CP>	<PREF_CH_NAME>	Contact Preferences	Preferred Contact Channel Name
Preferred channel ID	CDMETHODSTATUSTP	METHOD_ST_TP_CD	<CP>	<PREF_CH_ID>	Contact Preferences	Preferred Contact Channel Id

Figure 4-17   MDM to BDW column mappings

Figure 4-17 shows how attributes from the Customer Master Repository in MDM are mapped to attributes in BDW, and you already know how attributes are mapped to columns on either side. With this insight, you can create the table definitions for the BDW tables IP, IDV, PST_ADR, TEL_ADR, and EMAIL_ADR.

Finally, you can associate the individual columns on the source side with the CUSTOMER_MASTER stage, and the columns on the target side with the WAREHOUSE stage, as shown in Figure 4-15 on page 79.

With the completed source and target stage configuration, the remaining steps are to describe the mapping logic. This can be done in the transformer stage, as depicted in Figure 4-18.
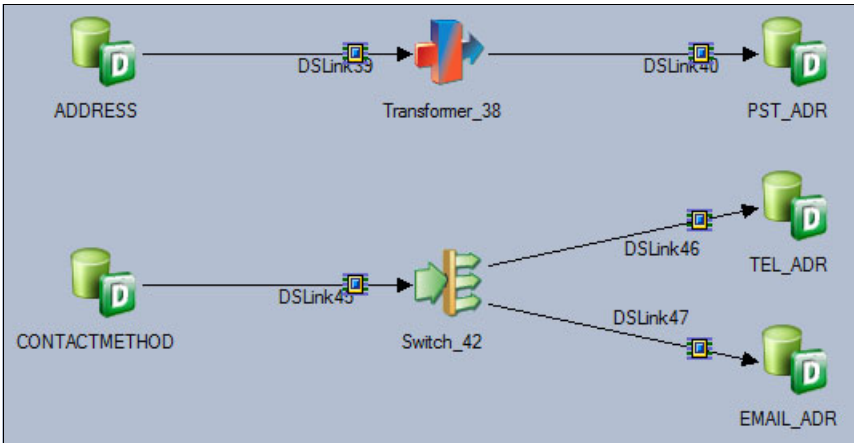
Figure 4-18   Transformer stage

The transformations can include scalar functions to manipulate individual records. A join, for example, is expressed with a join stage. DataStage includes a wide range of pre-built stages that a user can extend with an implementation of a user defined operator.

The transformations can then be implemented by using the Transformer Stage Mapping Editor, as depicted in Figure 4-19.
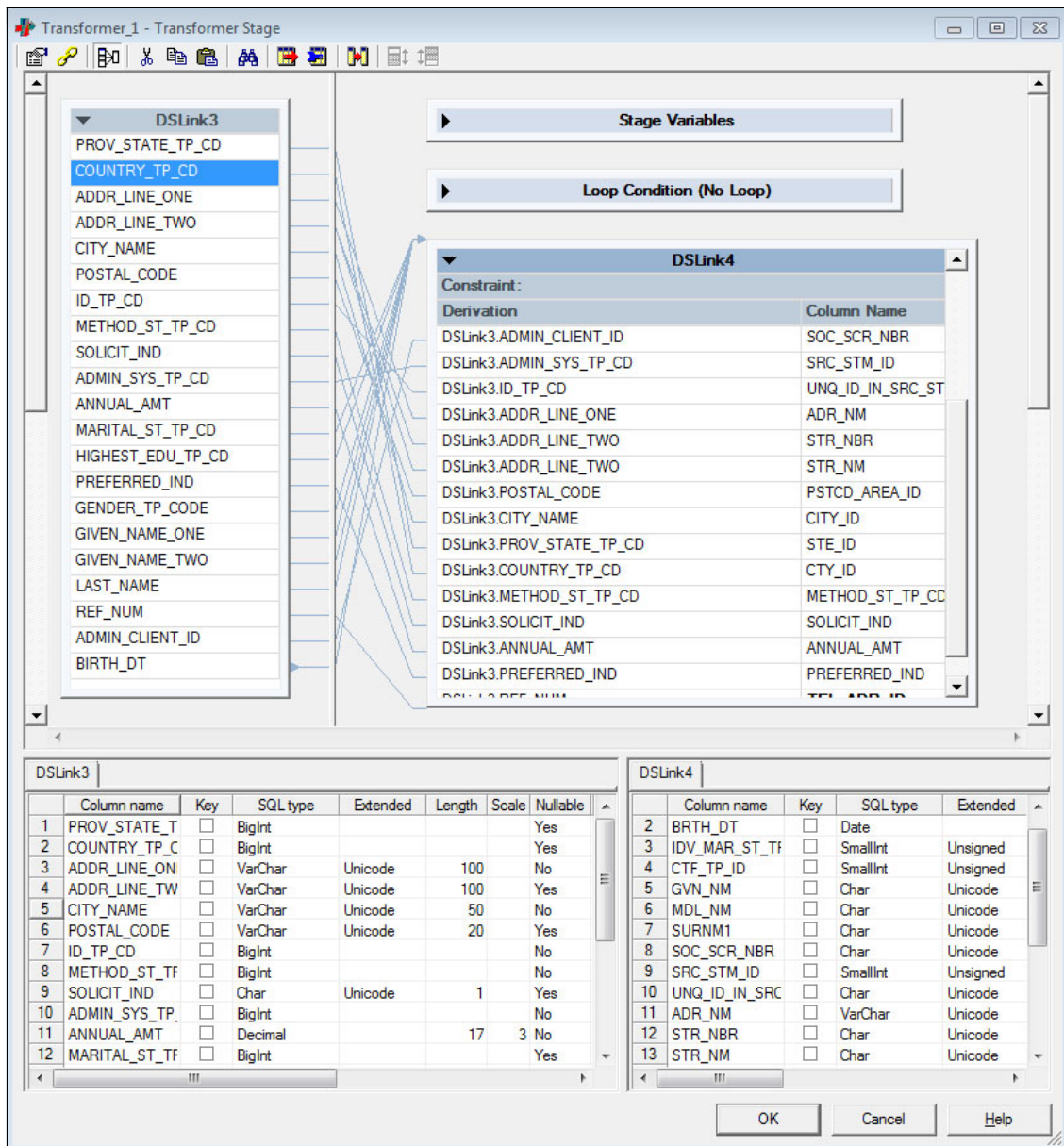

Figure 4-19   Transformer stage mapping editor

4.6.2 Executing the DataStage job

Before any DataStage job is executed, it must be compiled and deployed to the DataStage runtime engine. The details of DataStage job compilation are outlined in the InfoSphere Information Server documentation.

Upon execution, the job produces operational metadata that can be viewed in an operations console. It indicates how well the transformation worked and whether problems with individual records exist. The DataStage Operations Console is also described in detail in the InfoSphere Information Server documentation.

Any DataStage job can be configured to produce operational lineage metadata with the execution output, and thus connect the source Customer Master Repository with the target data warehouse in much the same way a Data Click activity can. A detailed description of lineage metadata is in Chapter 5, "From Master Data Management to PureData for Analytics" on page 85.

4.7 Summary

This chapter works with the architecture blueprint to complete the information supply chain from the Customer Master Repository to the marketing repository. Two main integration links in the blueprint and several techniques are available to implement these links.

The chapter describes Data Click as an implementation geared toward a self-service user, and DataStage ETL as a means to implement complex transformations. Both operations are driven by metadata; an explanation describes how physical and lineage metadata is made available to InfoSphere Information Server.

The chapter describes the attributes and mappings needed to build the marketing repository.

5

From Master Data Management to PureData for Analytics

This chapter describes the use of IBM Information Server to control the information supply chain from InfoSphere Master Data Management (MDM) to PureData for Analytics. It includes a description of all aspects of a managed information supply chain including governance and automation.

5.1 Governance

Governance is a complex practice concerned with the integrity of all information assets in an enterprise, including hardware, software, and data. It has been studied for several years and is being enforced in many enterprises at different levels of maturity. Our discussion of governance is concerned with the data governance along the information supply chain, as depicted in Figure 4-1 on page 59. We discuss governance aspects of stewardship, business definition, policy definitions, and data quality.

5.1.1 Stewardship

Stewards are specific personas responsible for the semantics and integrity of the metadata assets under their management. Stewardship therefore is the management of stewards and their asset assignments. Information Server administers stewards through user IDs.

An Information Server user does not have to become a steward, but a steward has to be an Information Server user with the appropriate privileges.

Having imported a full Customer Master Repository physical database schema, along with objects and services (in Chapter 4, "PureData for Analytics: The marketing data repository" on page 57), we can now associate a data steward to any metadata asset. Open the asset definition in Metadata Workbench and assign a steward.

Figure 5-1 is the listing of all assets managed by a user called Example Steward. We assigned four physical tables from the MDM schema to Example Steward and also three policy definitions. There are currently no Customer Master Repository objects or services under management by Example Steward.
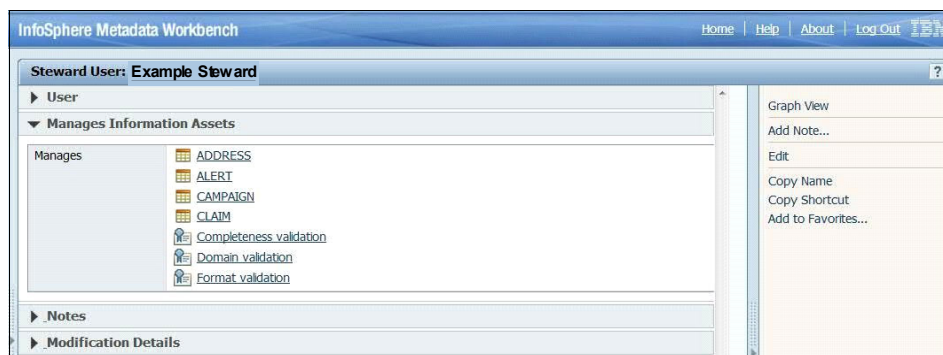

Figure 5-1   Managed information assets

With such an assignment, we can track stewardship of an asset throughout all tools in the Information Server suite and throughout all lifecycle phases of the asset. If, for example, a steward is assigned to a table, it will be present in the lineage graph including that table, in the business term definitions that exist for the table, in the policy definitions related to the table, or in the quality rules bound to the table. If quality exceptions are found for the table, they will automatically be owned by the steward for remediation. If the table definition is exported or migrated from one Information Server instance to another, it will retain the stewardship assignment and the definition of the steward. Metadata administrators are able to unassign or reassign stewards from their assets or delete the asset definition from the metadata repository.

5.1.2 Business

A typical challenge for data management is the difference in concerns users have with the data and the language used to describe those concerns. The marketer is interested in using the content of master data to run campaigns; whereas, technical users such as the MDM architect or a metadata administrator are concerned with, and responsible for, the design, quality, and integrity of the MDM environment.

IBM InfoSphere Business Glossary is a tool in the Information Server foundation tool suite that was explicitly designed to bridge that gap. It allows a business user to manage the MDM assets through business terminology rather than requiring the business user to become a metadata expert. This means the business user does not need to become an expert in the physical tables, objects, or services used by MDM, but rather they are able to use the business vocabulary that has been mapped to the Master Data implementation.

Most enterprises already have some form of a business dictionary available for use, and sometimes it is implemented as a formal glossary. Often however, it is a spreadsheet or Microsoft Word document. The IBM InfoSphere Business Glossary tool includes tools to facilitate the import of external business dictionaries into Information Server. Chapter 2, "Using IBM Industry Models for marketing" on page 13 demonstrates the importing of the business glossary made available with the IBM Industry Models. The one used for our marketing campaign features more than 10,000 terms specifically for the banking and financial markets industry.

At the center of a business dictionary is a business term. Business terms are organized into business categories. Business terms can have typed relationships to other business terms, so that one term can be a synonym for another term, or a term can be used to replace another term. Business categories are containers for business terms that can be deeply nested resulting in a hierarchical category tree. Stewards can be assigned to specific terms or entire categories. Business

terms and categories can be put under workflow management where a term definition must follow an approval process with different user roles involved.

To illustrate the workflow management capabilities in IBM InfoSphere Business Glossary in the context of a Customer Master Repository, we extend the glossary and demonstrate how to manually create it using the IBM InfoSphere Business Glossary UI. We create the following two terms, instead of business terms with an actual business meaning:

► Address Line One
► Address Line Two

Figure 5-2 shows the definition of the term Address Line Two in a business category MDMModel. We also omit the description here and let you apply the proposed solution to a proper business term definition.

Figure 5-2   Creating a new business term

The terms Address Line One and Address Line Two are initially put into a candidate state. Figure 5-3 shows how the business term can be assigned to a physical asset in the metadata repository. Similar to stewardship assignments, we can use business terms to classify any type of asset in the repository, including physical tables, objects, services or logical concepts such as policies or business rules. The association is bidirectional and can be created from the term that classifies an asset or the asset that is classified by a term. For this example, we use Address Line One to classify the column ADDR_LINE_ONE, and Address Line Two to classify column ADDR_LINE_TWO. Both terms remain in a draft state at this point.
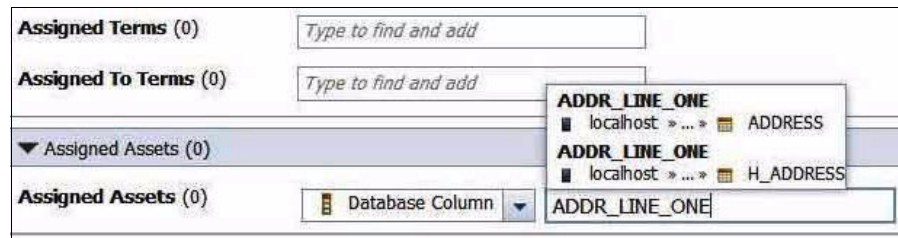

Figure 5-3   Completing a term definition

Figure 5-4 is used to request approval for only one of the two term definitions, Address Line One.
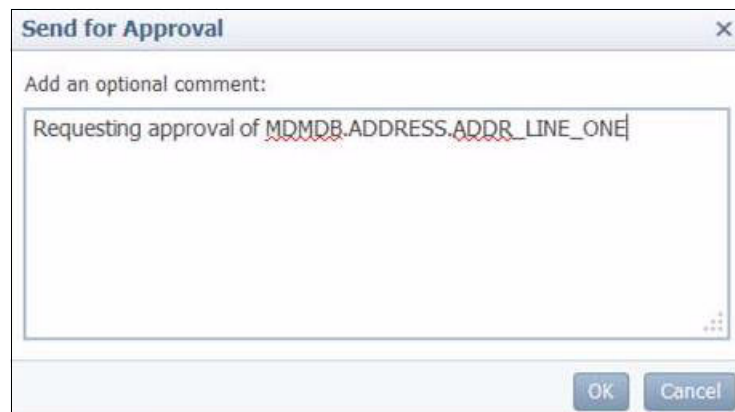
Figure 5-4   Requesting approval for a term

An approver sees the list of terms submitted for approval, as shown in Figure 5-5. Only Address Line One and *not* Address Line Two is shown in the listing. The approved term can then be published to the production repository, so it is available to any business user.
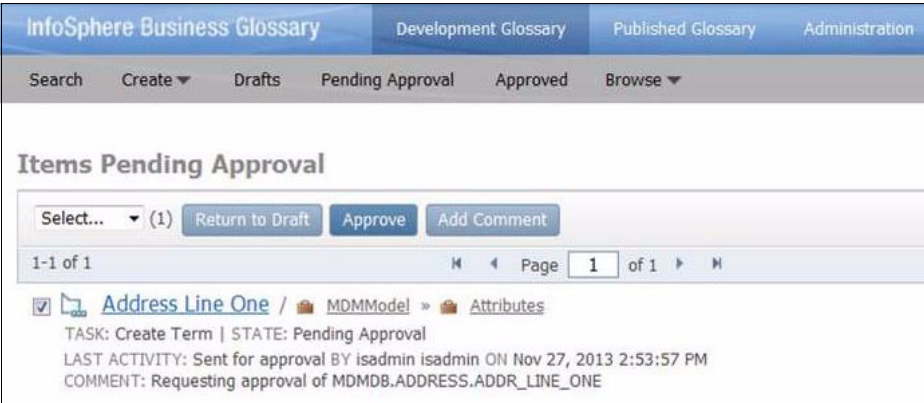
Figure 5-5   Items pending approval

Although we defined two terms and created classifications to two customer master repository columns, only the approved term is visible to the common business user using IBM InfoSphere Business Glossary. Figure 5-6 shows how "auto suggest" finds Address Line One with a search for `Address Line`. The business user has read access to Address Line One within the MDMModel category and the assets classified by the term.

Figure 5-6   Auto suggest for published business terms

Figure 5-7 depicts the approved business term and the classifications.
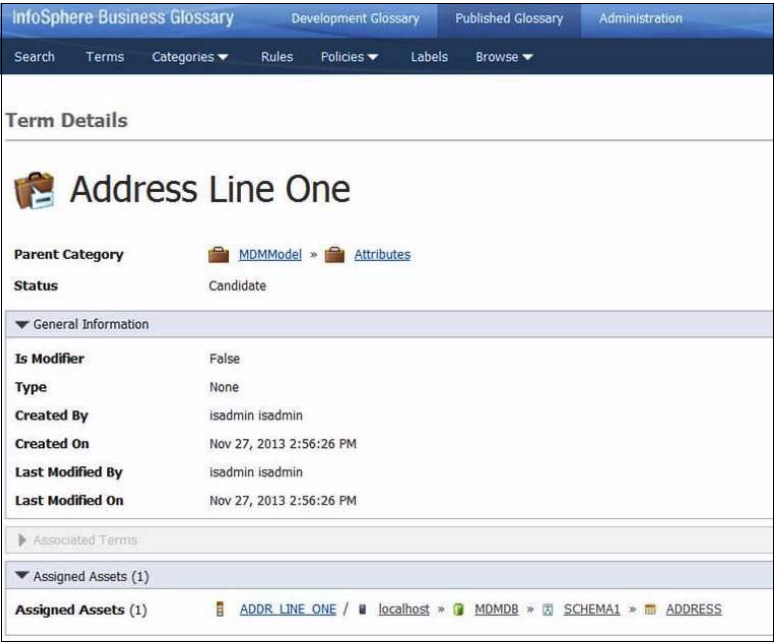
Figure 5-7   Approved business term

We introduce a business glossary as a means to achieve two goals:

► Allow a business user who is unfamiliar with technical entities to navigate the information supply chain, including the Customer Master Repository or any other entities introduced in Chapter 4, "PureData for Analytics: The marketing data repository" on page 57.

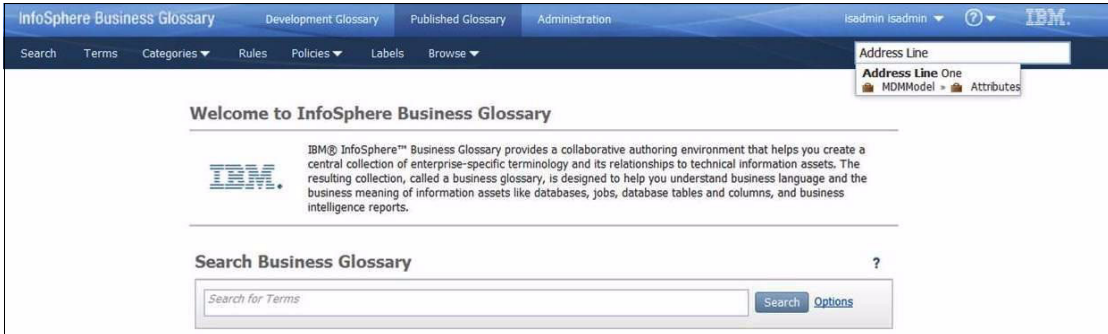► Introduce a means of access control and manage visibility to technical metadata through business terms. We leverage the workflow management of IBM InfoSphere Business Glossary to publish only approved terms that classify assets relevant for the business user.

Business terminology is again visible throughout the entire lifecycle of the classified asset and survives any export or migration. Business terms are a prominent entry point into much of our Information Server foundation tooling and not limited to IBM InfoSphere Business Glossary. Many search interfaces in the suite are designed with the business user and his terminology in mind.

5.1.3 Policy

IBM InfoSphere Business Glossary can be used to define data governance policies in addition to business vocabulary. Policy definitions are logical constructs that capture the intent of a data governance initiative but should be backed by data rules to validate the data against the policy definitions.

Policies can be freely defined by the business but are often demanded by regulators for specific industries. The verbal definition of these regulatory requirements can immediately be imported into IBM InfoSphere Business Glossary and form the basis of our policy-driven data governance. Policies are again organized in a tree structure much like business categories. Policies can have subpolicies, which in turn can have other subpolicies.

Figure 5-8 shows a sample policy tree with subpolicies to implement governance for the following items:

► Data compliance
► Data lifecycle management
► Data protection
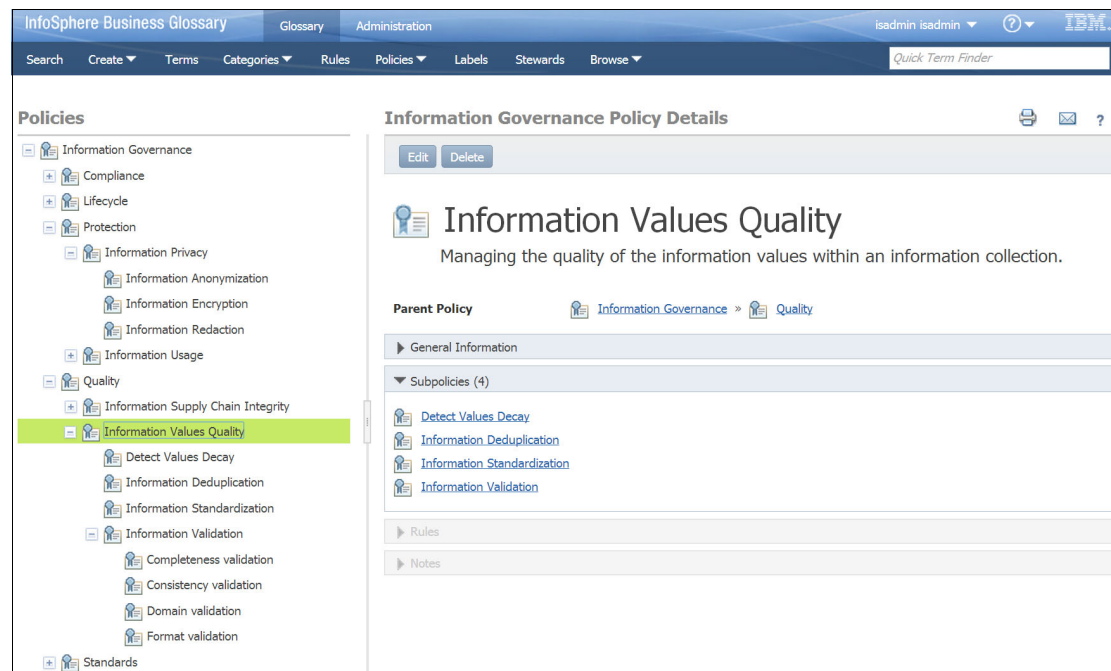► Data quality
► Data standardization

Figure 5-8   Information Governance policy tree

Data quality as a typical data governance initiative (see 5.1.4, "Quality" on page 94). Although the policy definitions are universally applicable to all data stores in our blueprint, we define data rules used to validate the consistency of the data in the data warehouse. That makes sense, especially if the data warehouse is fed by the Master Data Management system and other (non-managed) sources. The data warehouse is ultimately used to facilitate the marketing campaign, so we want to validate that the data we are submitting to the downstream systems is consistent according to the business requirements. See Figure 5-9. The same principles can again be applied to the Customer Master Repository or even the data sources.
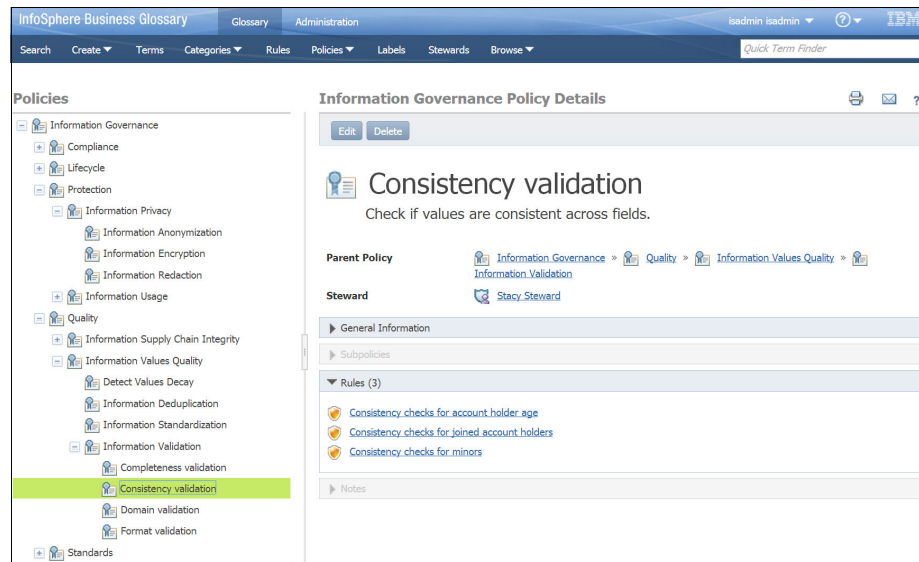
Figure 5-9   Information Governance policy tree

A policy definition is typically created and owned by a data steward. It has a textual description and is implemented through any number of governance rules. The data steward is the one who associates the governance rules with the policy definition.

Governance rules are still logical concepts only, but more atomic than a policy. A governance rule is used to describe a particular aspect of the policy. The following tasks are examples of rules:

► Validate that the street address information for a person is complete.
► Validate that the person age is greater than 18 and lesser than 125 years.
► Validate that a Social Security number has the expected format.
► Check whether joint account holders are married or cohabitant.

These rules are specific to the industry, the line of business, and data management initiative. A direct mail marketing campaign, for example, must have valid address data but a phone marketing campaign only needs valid phone numbers. Therefore, the number of rules can grow rather large but is easy to manage in the policy tree. IBM InfoSphere Business Glossary provides the same level of workflow management, and navigate and search functions that you have already seen for business terminology.

5.1.4 Quality

This section describes a data rule implementation for a single governance rule for our marketing campaign.

Consider the data rule in Figure 5-10 on page 95. It implements the governance rule to check whether joint account holders are registered as married or cohabitant individuals. The rule is expressed in syntax suitable for execution by IBM InfoSphere Information Analyzer.

The following source code expression is the result of a rule definition using a rule editor in IBM InfoSphere Information Analyzer.

```
if JOINED_ACCOUNT='YES' THEN MARITAL_STATUS in_reference_list
{'married', 'cohabitant'}
```

Data rule definitions generated with IBM InfoSphere Information Analyzer can be complex and contain variables, reference lists, multiple conditions, and complex expressions. The data rule is developed and bound to a set of physical columns. The columns can be from a single table or span multiple tables even across database boundaries. In fact, the same rule definition can be bound to multiple sources suitable for the data rule definition.

The rule can also be tied to a benchmark and does not necessarily have to be a 100% successful execution. Sometimes, we expect 50% of records to be male and 50% to be female. Any 5% variation from this benchmark might be considered a violation.

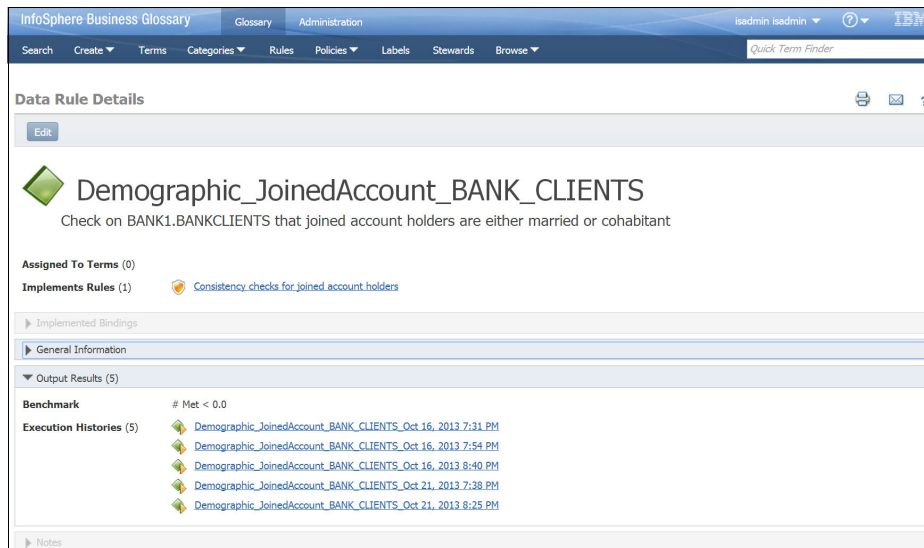In our rule definition, however, we set the benchmark as `# Met < 0.0`.

Figure 5-10   Data rule

The rule is compiled and executed by IBM InfoSphere Information Analyzer for either the entire data set or a data sample. It can be executed on demand or scheduled. As shown in Figure 5-10, with the IBM InfoSphere Business Glossary UI, we can inspect a high-level result for every rule execution directly from within the rule page.

This rule was created in IBM InfoSphere Information Analyzer but is available for inspection through IBM InfoSphere Business Glossary. For more detailed execution results including trend lines, benchmark comparisons and the list of violating records, you can use the IBM InfoSphere Information Analyzer desktop client.

With the help of data rules, you are able to validate the Customer Master Repository centric information supply chain and analyze the data at any point in the chain. The analysis is done for data that is in place and not for data that is in motion. It should not be confused with the quality logic that can be enforced at a transactional level where individual records are rejected if they violate a specific data rule. We do not perform the rule-based analysis because we doubt the integrity of the load processes (initial load and real time updates). Instead, we do the analysis through data rules so we can learn about the data itself and how well it meets the business requirements captured in the policies.

5.2 Automation

At this point we have achieved complete management and metadata integration for the Customer Master Repository and its role in the information supply chain, as depicted in Figure 4-1 on page 59. What we have not done yet is a continuous integration between the Customer Master Repository, the data warehouse, and the data marts. We already configured Data Click between the Customer Master Repository and the data warehouse in 4.3, "Data integration" on page 60.

We can use one or multiple blueprints to configure the Data Click Activities for scheduled execution or to continuously update the data warehouse with changes in the Customer Master Repository. We can also use a Data Click to push the changes from the data warehouse directly into the data marts hosted in the PureSystems for Analytics.

For scheduled execution we use the DataStage engine with Data Click, and no special configuration is required. Any DataStage job can be executed in a scheduled manner, including any job we created in 4.6, "DataStage" on page 78. DataStage provides a scheduler and Data Click makes a subset of the scheduling functions available through the Data Click UI.

For a real-time synchronization we configure the Data Click activity exactly as in section 4.5.1, "Preparing the Data Click execution" on page 73, but instead of DataStage the administrator uses a IBM InfoSphere Change Data Capture (CDC) engine. CDC is a replication offering as part of the InfoSphere brand but licensed separately. The CDC engine can be deployed on a remote server and Data Click has to be able to connect to the CDC server only.

CDC uses a publication-subscription model where the publisher makes any transaction on the source side known to its subscribers. So, for any transaction that inserts, updates, or deletes a record on the MDM data hub, CDC will capture the transaction from the transaction log and allow the subscribers to apply it identically on the target. The DBMS types used on the capture side and the apply side do not have to match. Data Click can use CDC to capture changes on the Data Warehouse and apply transactions onto the Data Marts.

The configuration for CDC is done by the administrator in Blueprint Director and described in the Data Click user documentation.

The deployed blueprint is accessible to the self-service user as described previously. The user again launches the Data Click configuration wizard in the blueprint and now finds the list of tables that already exist in the target, as shown in Figure 5-11.
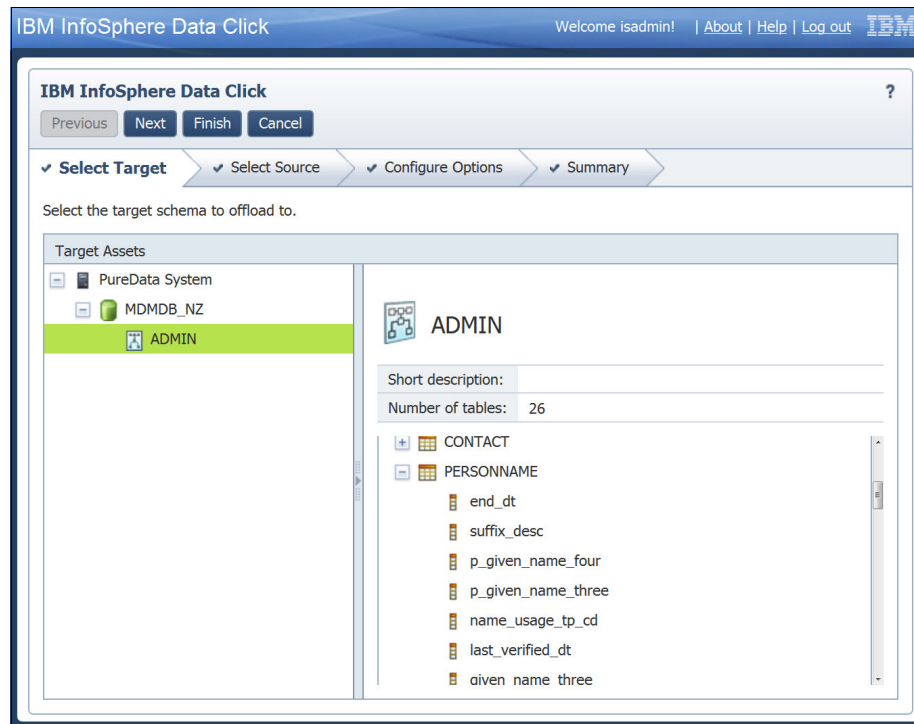
Figure 5-11   Data Click for continuous update: target page

On the source page, the user selects the columns and tables that this user wants to monitor for changes.

Note: The tables do not have to exist yet in the target environment, and CDC is able to create the tables and perform an initial load with the records from the source. CDC subsequently creates the publication and subscription needed to synchronize the source and target table.

On the Configuration Options page (Figure 5-12), the user selects **Continuously Update** with optional parameters for the target table generation.

Figure 5-12   Data Click for continuous update: configuration options

Figure 5-13 shows the Summary page where we find CDC automatically selected as engine for the request. Notice again that the self-service user does not need any CDC credentials or knowledge about the use of CDC in order to facilitate the replication.

Figure 5-13   Data Click for continues update: summary page

Data Click creates the subscription in CDC and puts it into an active and running state. The administrator can use the native tooling provided by CDC to monitor that status of a subscription or change it.

For the lifetime of the subscription, CDC will replicate the changes on the source tables in the Customer Master Repository to the target tables in the warehouse.

Complete lineage

Figure 5-14 and Figure 5-15 on page 101 are two parts that show the lineage graph now completed with the mapping generated by Data Click, based on the active subscription in CDC.

The two parts can be read from left to right; consider the following information:

1. The BL_060_AI_PersonName job has been used to load MDMDB PERSONNAME table.

2. The addPerson method loads the PERSONNAME table through the use of a TCRMPersonNameBObj.

3. A subscription has been created by user `holger` in project default on October 11, 2012 to synchronize the P_GIVEN_NAME_FOUR and P_LAST_NAME columns from the MDMDB with a MDMDB_NZ on a PureData system
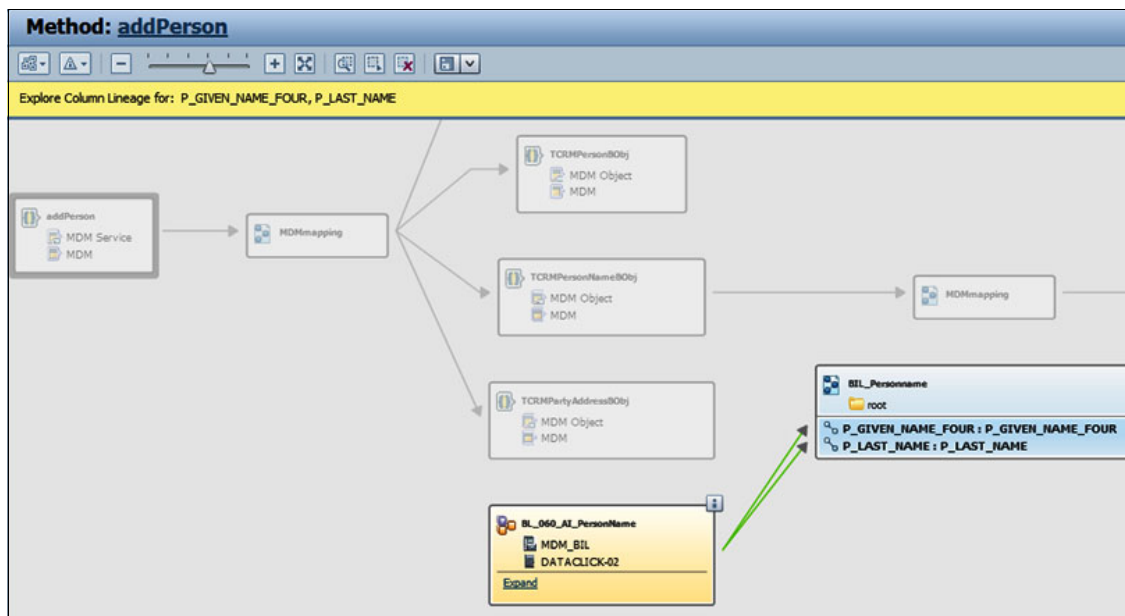
Figure 5-14   Complete end-to-end lineage graph with replication (part 1 of 2)
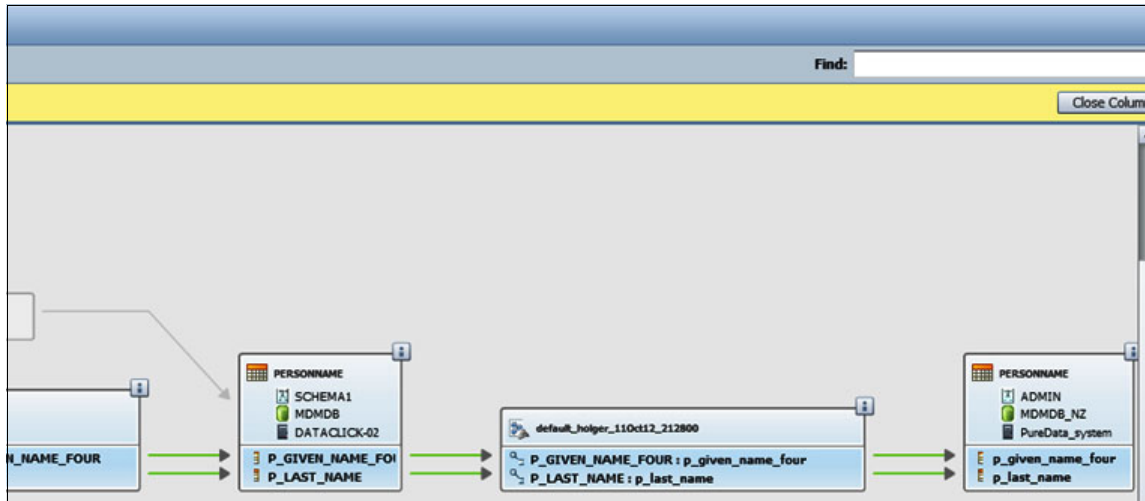
Figure 5-15   Complete end-to-end lineage graph with replication (part 2 of 2)

The metadata has been automatically constructed by Data Click on behalf of CDC and complements the metadata story covered in Chapter 4, "PureData for Analytics: The marketing data repository" on page 57 and Chapter 5, "From Master Data Management to PureData for Analytics" on page 85. This closes a loop of metadata integration with MDM where:

► Metadata has been used to describe and govern master data.
► Metadata has been used to operate over master data.
► Operations over master data generate new metadata for governance.

The lineage graph can be continued to include the EMM system shown in Figure 4-1 on page 59 on the consuming side, or the source systems on the producing side. The metadata administrator can choose to import as much metadata as is useful, and complete the lineage graphs accordingly.

5.3 Summary

This chapter continues the discussion of data governance and included stewardship, business definition, policies, and quality rules as applicable to the marketing campaign scenario. It shows how to use the workflow management capabilities in IBM InfoSphere Business Glossary and apply that to individual metadata assets from the Customer Master Repository. It also demonstrates how Data Click can be used to facilitate data synchronization between the Customer Master Repository and the data warehouse through the use of replication.

6

Leveraging Customer Data within IBM Campaign

This chapter describes several foundational components required for a working knowledge of IBM Campaign. The examples are in context with the same use case referred to throughout this book.

This chapter explains the reasons why marketers use IBM Campaign and how they use it to develop target marketing that can help to reduce the risk of customer attrition (*churn*). It also describes how marketers might use the data generated and managed by their data management technologies to support their efforts.

First, gaining a high-level working knowledge of the IBM Enterprise Marketing Management (EMM) Value Proposition as realized by IBM clients can be helpful to completely understand the value proposition of the IBM Campaign solution.

6.1 EMM Omni-Channel Customer Engagement

IBM EMM Omni-Channel Customer Engagement (OCCE) technology helps marketers target millions of individuals every day with the correct message at the correct time through the correct channel. Channels are messaging delivery or touch points used to communicate with customers. Examples include email, online stores, text messaging, and direct mail.

OCCE provides the tools to track customer actions while honoring individual customer preferences. It also provides the tools to quickly and cost-effectively design, execute, and measure customer-driven communication strategies across all channels, both online and offline. Customer actions can be recorded and used within campaign decision logic while respecting individual customer-directed preferences, such as opt-out and do-not-track.

By using IBM Campaign, marketers can accomplish the following actions:

► Build an effective, ongoing dialog with a customer or prospect through a complete interaction history, which can include offers presented, context and campaign details, and response or non-response.

► Manage the entirety of the campaign's logic and lifecycle, including audience segmentation, exclusions, and assignment of offers and channels.

► Enable use, reuse, and tracking of campaign building blocks, including audiences, segments, offers, treatments, and exclusion rules.

► Close the loop by attributing responses of various types (direct, indirect, inferred) to campaigns using several attribution methods (best match, fractional, multiple or custom).

► Integrate digital marketing and cross-channel marketing efforts by using a single software platform that allows both efforts to leverage the same data and marketing building blocks.

6.2 IBM Campaign: Introduction

IBM Campaign is a web-based Enterprise Marketing Management solution that enables marketers to design, run, and analyze direct marketing campaigns.

Marketing professionals typically use IBM Campaign in the following ways:

► Administrators perform initial and ongoing tasks such as adjusting configuration settings, mapping database tables, and defining custom attributes and offer templates.

► Marketing users create and run direct marketing campaigns.

To conduct a marketing campaign, users begin by defining offers for various target audiences. Then, they build a process, or flowchart, which provides a visual representation of the campaign logic. Part of building a flowchart involves associating offers with target audiences.

As users design campaigns, they access, merge, and manipulate data from multiple sources, including relational databases and flat files. In this case, marketers select customer data from the Marketing Repository, which contains core customer profile and transaction details, define suppression logic, segment the remaining customers, sample the data, assign offers, and produce the marketing output files.

IBM Campaign provides tools that enable marketers to query and manipulate data using an intuitive drag-and-drop interface. Marketers do not need to understand Structured Query Language (SQL) or engage IT resources to build their campaign logic.

The platform is also highly extensible to enable customers to build customized functions and queries through the support of raw SQL, macros, and functions.

After building a campaign flowchart and assigning offers to the appropriate target segments, users can easily run the flowchart to generate a contact list. Flowcharts can be run in test and production mode, enabling the marketer to run "what if" scenarios until they are satisfied the results will meet the campaign requirements. To control the timing of marketing campaigns, users can schedule campaigns to run at any required time, either once or on a periodic basis.

During the course of a campaign, contact history and response history are recorded. IBM Campaign uses this history to track and analyze campaign results, so users can refine their campaigns over time and measure the overall success and impact on the business.

IBM Campaign consists of a front-end web application server and back-end processing to handle flowchart run-requests. IBM Campaign runs on a common

foundation called the IBM Marketing Platform. That platform provides all IBM EMM modules with built-in roles-based security, authentication, and authorization.

6.3 IBM Campaign: Concepts

A campaign is defined by a business objective, a corporate-defined initiative that is specific to a marketing plan, and a date range during which the campaign is effective. Each campaign consists of one or more flowcharts, where users select recipients and assign offers.

Campaign flowcharts perform a series of actions with data to execute the campaign. For this example, users create a retention campaign to deliver an offer to customers who otherwise might be lost through attrition. This campaign consists of two flowcharts:

► One that generates lists of customers to whom the offer will be sent
► One to track responses to the offer, for reporting and analysis

6.3.1 Flowcharts

Within IBM Campaign, flowcharts represent a sequence of actions that are performed with relevant data, as defined by building blocks called *processes*.

Each campaign consists of at least one flowchart. Users configure and then connect processes in each flowchart to achieve specific marketing goals. For example, this flowchart selects the correct recipients for the "Customer Winback" campaign, which is focused on attrition risks, assigns various the appropriate offers to recipients, and then generates the outbound marketing lists. A second flowchart tracks respondents to the campaign, so users can calculate the impact of this new marketing initiative.

A sample flowchart is depicted in Figure 6-1.
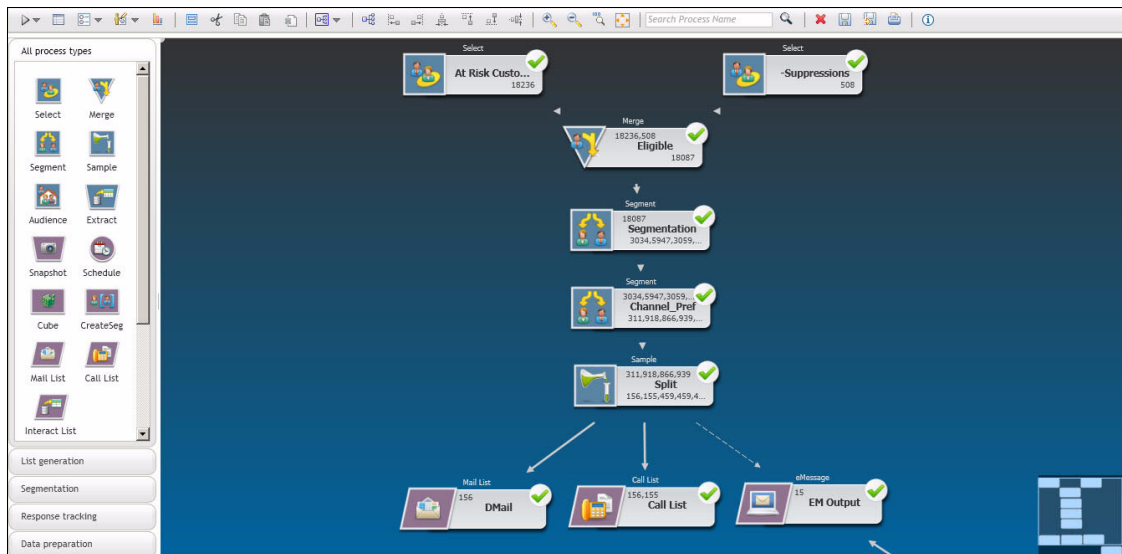
Figure 6-1   Simple flowchart example

Flowcharts can connect to one or more operational data sources. A data source contains information about the company's customers, prospects, or products, for use in marketing campaigns. In this example, a customer segment is being read along with an updated list of suppressions to guard against contacting customers who might have opted out of a particular communication channel.

Flowcharts can be run manually, scheduled, or initiated by an outside process or trigger.

Note: Licensed users of IBM Interact can use IBM Campaign to run real-time interactive flowcharts that depend on the occurrence of an event. For more information about interactive flowcharts, see the *IBM Interact User's Guide*.

6.3.2 Processes

Flowcharts consist of individual processes that are configured to perform a particular task in a campaign or session, such as selecting data, merging two distinct audience groups, or writing the results of the campaign.

6.3.3 Offers

An *offer* represents a single marketing message, which can be delivered in a variety of ways. An offer can be used in one or more marketing campaigns.

Offers are reusable in these situations:

- ▶ In different campaigns
- ▶ At different points in time
- ▶ For different groups of people (cells)
- ▶ As different versions (by varying the offer's parameterized attributes)

Marketers can assign offers to target cells in flowcharts using one of the contact processes, such as Mail List or Call List, and can track the campaign results by capturing data about customers who received the offer and those who responded. In this example, the offers are designed to mitigate attrition risk among the customer base.

6.3.4 Cells

A *cell* is a collection of audience identifiers (such as customer or prospect ID) from the Marketing Repository. Marketers create cells by configuring and running data manipulation processes in Campaign flowcharts.

These output cells can also be used as input for other processes in the same flowchart (downstream from the process that created them).

Cells to which marketers assign one or more offers are called *target cells*. A target cell is a group of homogeneous individuals (or whatever the entity is for which the audience level is defined, such as individual customers or household accounts).

In this example, cells are created for high-value customers as defined by specific value segments, created by the marketer using IBM Campaign, who have not been active in the last several weeks and have opted in to receive both email and direct mail communications. Each cell created can be treated differently, receiving different offers or communications through different channels.

Cells that contain identifiers who are qualified to receive an offer but who are excluded from receiving the offer for analysis purposes are called *control cells*. In Campaign, controls are always hold-out controls.

6.4 IBM Campaign: Process

Before meaningful work can be done in IBM Campaign, initial configuration is required. Database tables must be mapped, data objects such as segments, dimensions, or cubes might need to be created, and individual campaigns must be planned and designed.

For detailed information about initial and ongoing configuration and administration, see the *IBM Campaign Administrator's Guide*. All referenced documentation is at the following location:

http://www-947.ibm.com/support/entry/portal/overview

6.4.1 Before creating campaigns

Before marketers begin using IBM Campaign to create marketing campaigns, there are some important initial tasks. Some of these initial tasks, such as creating offer templates, typically are done by administrators.

One of the most important initial tasks is to make information about customers and products available to Campaign. To make the marketing data available for use in Campaign, the database tables and files must be mapped. An administrator typically performs this step. The basic process is discussed in 6.4.3, "Mapping to customer and other data for IBM campaign" on page 112. An administrator also creates offer templates, strategic segments, and other data objects for use in campaigns. For detailed information, see the *IBM Campaign Administrator's Guide*.

After the initial objects are created and tables are mapped, marketers can begin creating marketing campaigns.

6.4.2 Example: A Customer Winback campaign

This example shows two flowcharts that were designed for the marketing campaign based on attrition risk (*retention*).

The retention campaign uses multiple channels to deliver an offer to customers who otherwise might be lost through attrition. The following two flowcharts are used:

► A contact flowchart generates lists of customers to whom the offer will be sent, using a different channel for each segment.

► A response flowchart tracks the responses to the offer and makes the response data available for reporting and analysis.

Contact flowchart

The flowchart in Figure 6-2 shows the sample retention campaign contact flowchart. This flowchart selects the eligible customers in each value segment and outputs contact lists for a different channel for each segment.
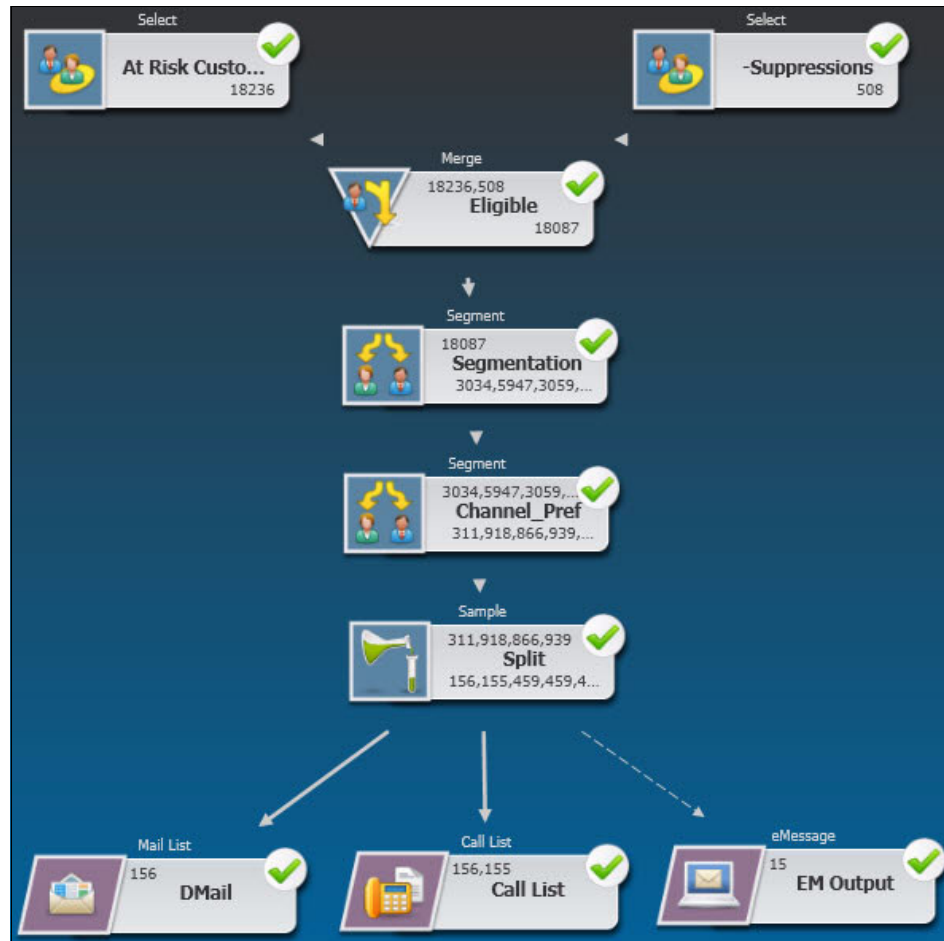

Figure 6-2   Retention campaign flowchart example

At the first level of the flowchart, the Select processes are used to select customers exhibiting the specified characteristics of attrition risk, and also customers who opted out of marketing communications across channels.

At the second level, a Merge process combines the "at risk" customers and excludes customers who have opted out.

At the third level, a Segment process divides all eligible customers into value tiers based on marketer defined criteria.

At the fourth level, each customer is assigned to a list. The high-value customers are output to a call list, so they can be contacted with the offer by phone. The medium-value customers are output to a mail list so they will receive a direct mail offer. The lowest value customers will receive the offer by email.

Response flowchart

A second flowchart in this same campaign tracks responses to the direct mail, email, and telephone offers captured by the call center and response capture systems. The response information is compiled outside of the Campaign application. For example, the call center might record responses in the database or a separate flat file. When the response information is made available to Campaign, the response flowchart can then query the data.

Figure 6-3 shows the retention campaign's response tracking flowchart. The Response process box evaluates which responses are considered valid, and how they are credited back to campaigns or offers. The output of the Response process is written to several response history system tables, where the data can be accessed for analysis using Campaign performance and profitability reports.

Figure 6-3   Response flowchart example

6.4.3 Mapping to customer and other data for IBM campaign

This section describes the process of mapping to customer and other marketing data for use in IBM Campaign. Although many other scenarios and methods for using marketing-related data exist, the scope of this document does not permit describing them all. For more detailed information, see the IBM Campaign installation and administration guides.

IBM Campaign uses Universal Dynamic Interconnect (UDI) to create a metadata mapping layer that defines a base, dimension, or general table to the system. The data source name is mapped along with the tables and column names that the business wants to make available to the marketer. If administrators want to, they can define aliases for table column names that are "marketing-friendly" so that the use of the data is more intuitive.

The purpose of this data mapping approach is to provide business user nomenclature and data structures familiar to the marketing users without a need for deeper database technical skills. IBM Campaign administrators are typically more technically adept and might be part of either the marketing or IT organization.

Prerequisites

To map IBM Campaign to the customer data housed within the Marketing Repository, several prerequisites must be completed:

1. Complete installation of the EMM Platform as outlined in the Platform Installation Manual.

2. Definition of the IBM Campaign database.

3. Creation of an ODBC connection definition to the IBM Campaign database.

4. Configuration within IBM Campaign to connect the data sources to the ODBC definition.

Mapping to data sources from the Campaign Settings menu

The following steps map a new base record table so that data can become accessible to processes in flowcharts:

1. Select **Settings** → **Campaign Settings**. The Campaign Settings page opens. Click **Manage Table Mappings**. The Table Mappings window opens, as shown in Figure 6-4.
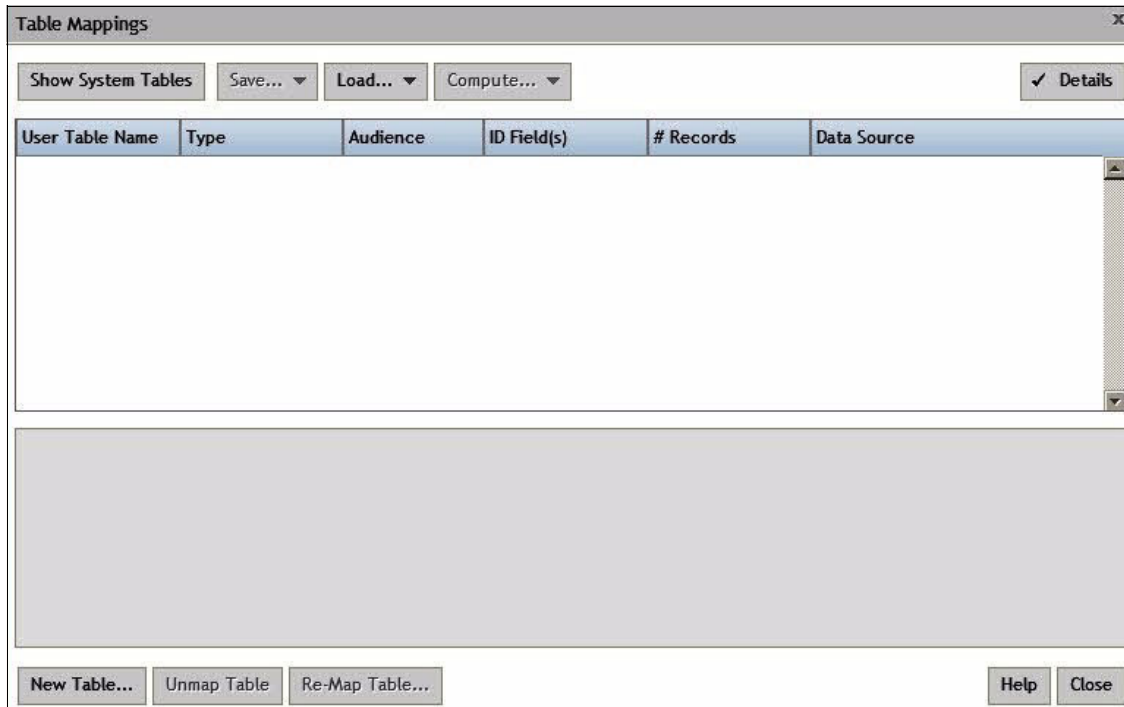
Figure 6-4   Table Mappings

2. Click **New Table Definition**, which takes you to the table definition input window. Then, select the table type window.

> **Note:** When users access the table mapping wizard from a Select process, the Dimension Table and General Table options are not listed.

3. If necessary, select **Base Record Table**, then click **Next**.

4. Select **Map to Existing Table** in Selected Database, select the data source name, then click **Next**.

5. Select the a table for mapping from the Source Table list.

   The tables are listed in alphabetical order by `<owner>.<table_name>`. If the expected table is not shown, ensure that the proper data source is configured to filter specific table entries.

   The source fields in the table that is selected are mapped automatically to new table fields in the base record table that is being created. To change the automatic mappings, select **fields** from the Source Table Fields list or the New Table Fields List, and use the **Add**, **Remove**, **Up 1**, and **Down 1** buttons, until the table is mapped as required.

   Clicking the **Field Name** column in the New Table Fields section automatically sorts column names alphabetically in ascending (or descending) order.

6. Click **Next**.

7. Optional: You may change the name that Campaign uses for the base record table and its fields to a more "user-friendly" value:

   – To change the table name, edit the name in the IBM Campaign Table Name field.

   – To change a field name, select the field name in the New Table Fields list, and edit the text in the IBM Campaign Field Name field.

   Click **Next**.

8. Select a level from the Audience Level drop-down list shown in Figure 6-5 on page 115. The Audience Fields list is automatically completed with the fields required to define the selected audience level. Match one or more fields in the new base table corresponding to each required key.

9. If the selected audience level is normalized (if each unique audience ID appears no more than once in the current base table), select the **This audience level is normalized in this table** check box.

   Setting this option correctly is important for correct configuration of options in an Audience process. If you do not know the correct setting, leave this option unselected.

10. Click **Next**.

11. Optional: In the Specify Additional Audience Levels window, specify one or more additional audience levels that are contained in the base record table.

12. Add more audience levels so users can use this table as a "switching table" to translate from one audience level to another using the Audience process in a flowchart. Complete the following steps

    a. Click **Add**. The audience level and ID fields window opens (Figure 6-5 on page 115).

    b. Select an Audience Level.

c. For each Audience Field, match appropriate field or fields from the base table to the corresponding key or keys in the audience level.

d. If the selected audience level is normalized (that is, each unique audience ID appears at most once in the current base table), select the **This audience level is normalized in this table** check box.
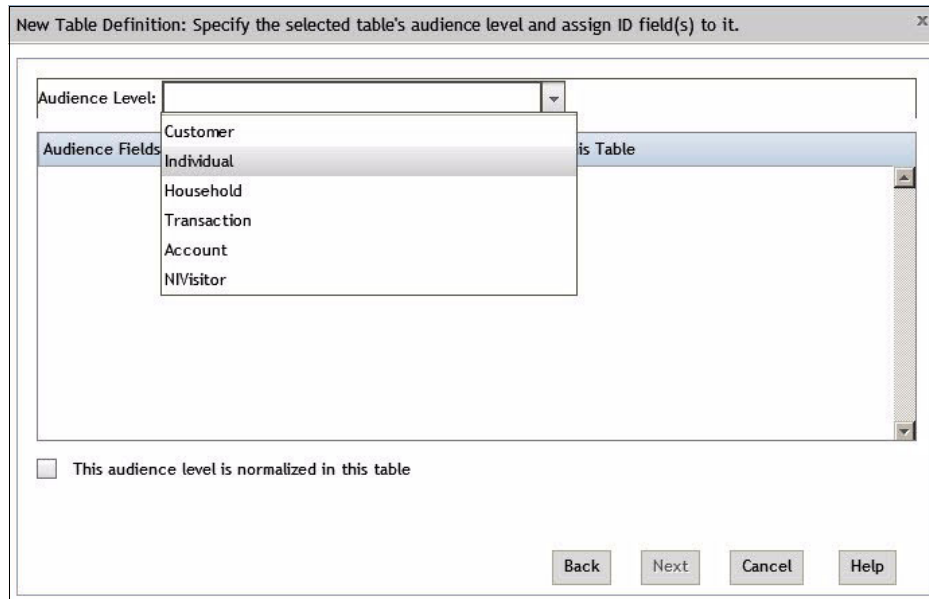

Figure 6-5   Selecting the audience level

e. Click **OK**.

f. Repeat step 12 on page 114 (steps a - f) for each audience level desired, to add for the base table, then click Next.

13. If dimension tables exist in the current table catalog, the Specify Relationship to Dimension Tables window opens; complete the following steps:

a. Select the box to the left of any dimension table that is related to the base record table being creating.

b. For each related dimension table, in the New Table's Key Field(s) To Use list, select a field from the base table to match each listed key in the Dimension Table's Key Field(s) list, then click **Next**.

14. Choose to precompute the distinct values and frequency counts for specific fields, or allow users to perform real-time profiling of data in the base record table.

15. Click **Finish**.

This completes mapping the base record table based on an existing database table. The new base table is part of the current table catalog, which you can manage through the table manager.

Accessing data sources from within a flowchart

IBM Campaign users, if given access permissions, also have the ability to map data for use in flowcharts, for use in several types of processes.

To access the customer or prospect database tables from within a flowchart, be sure that you are logged in to the referenced databases, as follows:

1. While editing a flowchart, click the **Admin** icon and select **Database Sources**, shown in Figure 6-6. The Database Sources window opens. The database that contains system tables is listed, and also all databases that Campaign is configured to access.
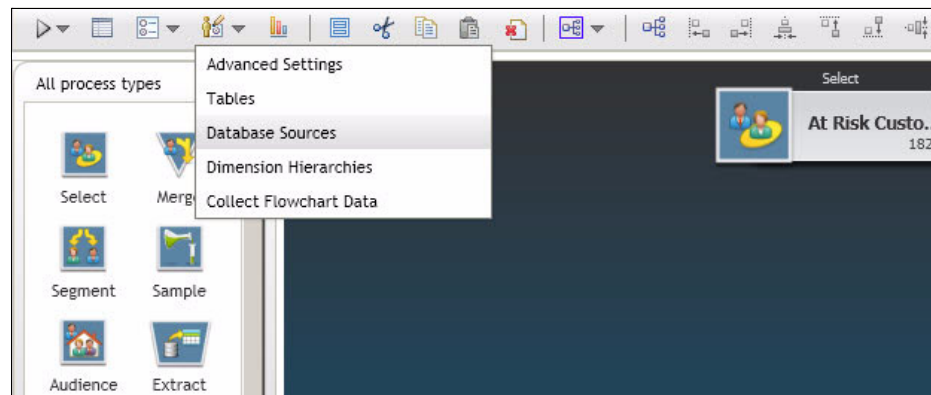

Figure 6-6   Table mapping from a flowchart

2. To log in to a database, select it and click **Login**.

3. Click **Close**.

The tables in the database are now accessible. To query a table in the database, first map the table, as described in the *IBM Campaign User's Guide*.

6.4.4 Defining campaign logic and configuring processes in flowcharts

IBM Campaign uses flowcharts to define campaign logic. Each flowchart in a campaign performs a sequence of actions on data that is stored in the Marketing Repository.

Each marketing campaign consists of at least one flowchart. Each flowchart consists of at least one process. First configure and then connect processes to perform data manipulation, contact list creation, or contact and response tracking for the campaign. For example, a Select process can select existing customers from the database by using defined criteria, and a Segment process can segment those customers into tiers.

By connecting a series of processes in a flowchart, then running that flowchart, campaigns are defined and implemented.

Campaign uses processes in a flowchart to perform various actions on the customer data. By adding process boxes to a flowchart, and configuring and connecting the processes, the logic for the marketing campaign is determined.

Most processes determine how Campaign selects and manipulates IDs from the user data. For example, first configure a Select process to identify high-value prospects, configure another Select process to identify medium-value prospects, then use a Merge process to combine the two lists. Configuring a process determines what occurs when the process runs.

Many types of processes are available, as Figure 6-7 on page 118 shows, and each one performs a distinct function. Some processes are intended to be used when implementing the campaign. For example, the Call List process assigns offers and generates a call list, which can be sent to a call center.

Other processes are used after the campaign is deployed. For example, use the Track process to update Contact History after deploying the campaign and see who responds. The processes are visible on the flowchart process palette at the left of the workspace.
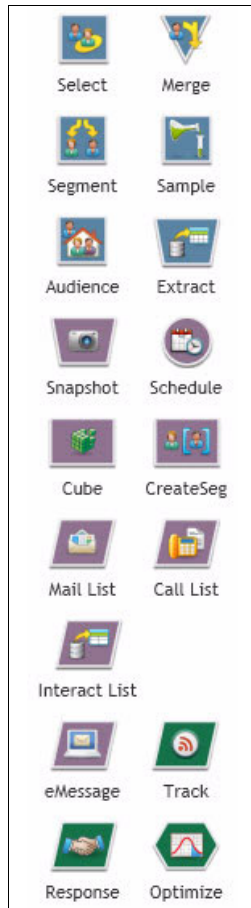
Figure 6-7   Process objects

To create a campaign flowchart, drag process boxes from the palette to the workspace. Then, configure each process box to do a specific operation, such as selecting customers to target for a mailing. Dragging connector lines from one box to another and connecting processes in the workspace determines the logical flow of events.

6.4.5 Creating queries to identify contacts

Queries are used to select, segment, or extract data from the data sources when designing a marketing campaign in IBM Campaign.

When configuring a Select, Segment, or Extract process in a flowchart, use a query to identify contacts from the databases or flat files. In our example

(Figure 6-8), the marketing user will select customers with high value and lapsed activity to target them for the "Customer Winback" retention campaign.

Figure 6-8   Selection process, accessing Marketing Repository data

6.4.6 Creating and managing offers

Offers are specific marketing communications sent to particular groups of people, using one or more channels. Each offer is based on an offer template defined by a campaign administrator.

In IBM Campaign, offers can be used in one or more campaigns; offers are reusable in these situations:

- ► In different campaigns
- ► At different points in time
- ► For different groups of people (cells)
- ► As different versions by varying the offer's parameterized fields

The general workflow is as follows:

1. Optional: An administrator defines custom attributes.
2. An administrator creates offer templates (required) and adds custom attributes to them (optional).
3. A user creates offers based on the templates.
4. A flowchart designer assigns the offers by configuring a contact process in a flowchart or by relating offers to target cells that were defined in the target cell spreadsheet.
5. The campaign runs, and the offer is made to the customer.

6.4.7 Assigning offers to cells in a flowchart

Flowchart designers assign offers to cells by configuring a contact process in a flowchart. A *cell* is a list of customers targeted to be contacted with a specific offer. A *target cell* is a cell that has an offer assigned to it. Optionally, control groups can be excluded from contact for analysis purposes.

Before beginning, create an offer so it is available for assignment. Also associate the offers with campaigns by using the Campaign Summary tab. Offers that are associated with a campaign are listed at the top of selection lists as "relevant" offers.

Note: Organizations that use a "top-down" management approach assign offers to cells in a target cell spreadsheet (TCS). A flowchart designer then selects recipients for those offers. For more information, see "Assigning offers to cells in a TCS" in the IBM Campaign Users Guide.

Complete the following steps to assign offers to cells in a flowchart:

1. Open a campaign and click a flowchart tab.
2. Click **Edit** in the flowchart window.
3. Double-click the contact process, such as Mail List or Call List, which contains the recipients for the offer.
4. Use the Treatment page (Figure 6-8 on page 119) to assign at least one offer to each cell.
5. Click the **Parameters** tab to see the names and values of each parameterized offer that was assigned on the Treatment page and to adjust the offer parameter values.
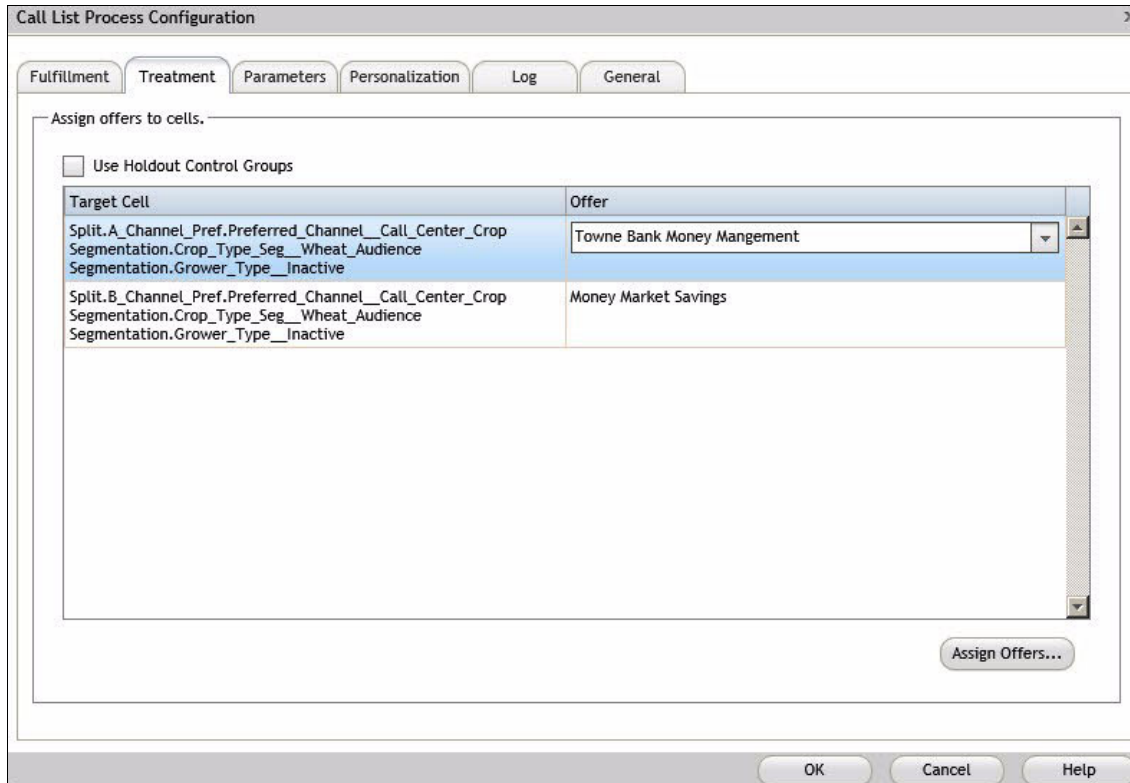
Figure 6-9   Assigning offers

6.4.8 Maintaining contact history

IBM Campaign maintains contact history to record information about offers that are sent to contacts and also information about control cells (for hold-outs or no-contacts).

The term *contact history* refers to information that Campaign maintains:

► What offers were sent
► To which customers (or accounts or households, depending on audience level)
► By which channel
► On what date

For example, a list of target customers can be generated as the output of a Call List or Mail List process in a campaign flowchart. Each target customer belongs to a cell that is assigned one or more offers. When the Call List or Mail List process is run in production mode with logging to contact history enabled, details are written to several tables in the Campaign system database.

Together, those tables comprise the contact history. The contact history records the specific offer version (including the values of parameterized offer attributes) given to each ID in each cell at flowchart run time. Contact history also records the members of control cells, who are purposely withheld from receiving any communications. Control cells indicate hold-out or no-contact controls, therefore customers belonging to control cells are not assigned any offers and are not included in contact process output lists.

6.4.9 Contact history and audience levels

IBM Campaign maintains contact history for each audience level, such as Customer and Household. Contact history provides a historical record of the direct marketing efforts, including who was contacted, what offers were made, by which channels.

Campaign maintains contact history in the system database tables, as follows:

► Base contact history (UA_ContactHistory) is recorded when all members of a cell are treated the same (when they are all given the same version of an offer).

► Detailed contact history (UA_DtlContactHist) is recorded only when individuals in the same cell receive different offer versions (offers with different values for personalized offer attributes) or a different number of offers.

   Detailed contact history can quickly grow very large, but it provides complete data to support detailed response tracking and to analyze targets and controls.

► For every production run, data is recorded in the Treatment tables (UA_Treatment). Control information for hold-outs is also recorded here.

   Hold-outs do not receive communications, but are measured against the target group for comparison. Treatment history is used together with contact history to form a complete historical record of offers that were sent.

Contact history and the corresponding response history are maintained for each audience level. For example, assume the campaign using two audience levels, Customer and Household. The table implementation depends on how the database is configured:

► Each audience level usually has its own set of contact and response history tables in the Campaign system database. That means the Customer audience level has a set of tables (contact history, detailed contact history, response history), and the Household audience level has its own set of tables.

► If the database is configured so that multiple audience levels write to the same underlying physical tables, then each audience level does not require its own set of tables. However, the underlying physical tables (contact history, detailed contact history, response history), must include a key for each audience level.

6.4.10 Tracking responses to campaigns

Use the Response process in a flowchart to track the actions that occur after a campaign. When running the Response process, data is recorded in the response history tables and is available for the Campaign performance reports.

Response tracking helps evaluate the effectiveness of the campaigns. Marketers can determine whether actions taken by individuals are in response to offers that were given. Users can evaluate responders and non-responders who were sent offers, and can also evaluate controls (individuals who were not sent offers) to see whether they performed the desired action despite not being contacted.

Campaign saves the response history and uses it in the Campaign performance reports, so you can easily determine the following information:

► Who responded: This is the list of audience entities (such as individual customers or households) whose behavior matched the response types being tracked.

► What they did, and when: IBM Campaign records the actions that were performed and the date and time of the actions. Examples are a click-through on a website or the purchase of a specific item. This information depends on either or both the Response Types that are set up and additional data that is captured during response processing.

► Which offer treatment they responded to: Any codes generated by Campaign (campaign, offer, cell, or treatment code) and any offer attributes with non-null values returned by the respondent are matched for response tracking.

► How the response is attributed: Criteria include matching codes generated by Campaign or non-null values for offer attributes, whether responders were in

the original targeted group or a control group, and whether the response was received before the expiration date.

► Additional information: The response history tables also record the following information:

– Whether the response was direct (one or more Campaign-generated codes were returned) or inferred (no response codes were returned).

– Whether the respondent was in a target cell or control cell.

– Whether the response was unique or a duplicate.

– Best, fractional, and multiple attribution scores.

– The response type (action) attributed to the response.

– Whether the response was received before or after the expiration date of the specific offer version. This information depends on the following property:

```
Settings > Configuration > Campaign > partitions > partition[n] >
server > flowchartConfig > AllowResponseNDaysAfterExpiration
```

The default value is 90 days.

Response attribution

Campaign supports three methods by which responses are credited to offers:

► Best match
► Fractional match
► Multiple match

All three of these response attribution methods are used simultaneously and recorded as part of response history. You can choose to use one, a combination, or all of these in the various performance reports to evaluate the campaign and offer performance.

Response attribution is performed on target treatment instances that remain after invalid responses are discarded (either because the treatment instance did not contact the responder, or because the target instance is expired).

For example, a respondent in a target cell that was given three offers returns a cell code; an exact treatment instance cannot be identified.

► Best-match attribution chooses one of the three offers to receive full credit.

► Fractional match attribution gives each of the three offers one-third credit each

► Multiple match attribution gives all three offers full credit for the response.

6.4.11 Reports

IBM Campaign reports can help you gather and analyze information about the campaigns. After running a campaign, marketers can study the results and then fine-tune the campaign strategy.

IBM Campaign reports analyze the campaign response rates, revenue, profit per responder, and other data. Based on this information, IBM Campaign can calculate the total and incremental revenue and profit, and the overall return on investment (ROI).

6.5 Summary

This chapter introduces the process marketers follow in using the most accurate customer data to produce an outbound, multi-channel campaign that is designed to reduce attrition risk for a specific group of customers.

Chapter 7, "Leveraging Customer Data within IBM Interact" on page 127 describes how marketers can extend this campaign to include inbound, real-time marketing communications.

7

Leveraging Customer Data within IBM Interact

This chapter describes several e foundational components that are required for a working knowledge of IBM Interact. The examples are in context with the banking use case referred to throughout this book. For specific use within this chapter, the use case is further defined, from the real-time marketer's perspective. The marketer has defined a specific real-time component to the outbound campaign described in Chapter 6, "Leveraging Customer Data within IBM Campaign" on page 103, and will use IBM Interact to target the appropriate customers with real-time marketing messages on the Bank website being used for our fictitious use case.

This chapter explains the reasons why marketers use IBM Interact and how the marketers use it to develop real-time marketing strategies, which are designed to reduce the risk of customer attrition (churn) through a "winback campaign." This chapter also covers how marketers can use the data generated and managed by the previously described data management technologies to support their efforts.

First, gaining a high-level working knowledge of the IBM Interact Value Proposition as realized by IBM clients is helpful in order to completely understand the functional application of the content in this chapter.

7.1 IBM Interact overview

IBM Interact is an interactive engine designed to deliver personalized offers to visitors of inbound marketing channels. Marketers can configure IBM Interact to be behavioral, strictly event-based, situational, strategic, and so forth. Interact is a module within the IBM suite of Enterprise Marketing Management tools, and is integrated with the IBM cross-channel marketing platform.

Interact uses both online and offline data, including in-depth historical knowledge of customers plus current customer activities, to create real-time customer interactions. These interactions can increase sales, build relationships, generate leads, increase conversion rates, optimize channel usage, and lower attrition. Marketers can establish the business rules and sales or service strategies that drive real-time interactions, and modify them to reflect the continuous learning from the marketing efforts.

Fully integrated with the IBM cross-channel marketing suite, IBM Interact gives your marketing organization the ability to coordinate real-time inbound customer treatment strategies with your traditional outbound campaigns. With Interact, marketers leverage the power of the IBM cross-channel marketing suite to enhance your real-time marketing efforts in several ways:

For example, you can use multi-channel operations in the following ways:

► Use information from customer touchpoints, which include web, call center, in-store, branch, and so on, to develop a rich understanding of customers and prospects, create a consistent brand, and maximize customer communications. See 7.4.10, "Touchpoints" on page 140.

► Create leading-edge, digitally delivered personalization.

► Engage known and anonymous visitors, consider behavior, and personalize interactions with them by making cross-sell offers, selecting editorial content, offering appropriate service options, and coordinating banner messages.

With Interact you can also optimize contact center interactions. For example you can leverage the power of your contact center for increased revenue generation and customer retention. You can provide smarter interactive voice response (IVR) navigation, on-hold message selection, instant offers for retention and cross-selling, prioritization of offers, and website intervention (such as with chat or Voice over Internet Protocol, VOIP).

IBM Interact enables marketers to control and fine-tune real-time marketing content, which is delivered to the relevant touchpoint systems. Strategies can include any factors considered important. These strategies can drive the response to specific customer actions or attrition risk, driving personalized

content from an instant-offer link on a website, to a cross-sell opportunity at a call center. IBM Interact provides control over critical online selling, marketing, and service strategies, and the ability to respond quickly to opportunities or changes in the marketplace.

7.2 IBM Interact: Baseline understanding

IBM Interact integrates with customer-facing systems, such as websites and call centers, and allows users to retrieve optimal personalized offers and visitor profile information in real-time to enrich the interactive customer experience.

Users configure Interact to integrate with customer touchpoints using an application programming interface (API). Using this API, users instruct Interact to gather a customer's information, add data to that information, and present offers based both on actions taken by the customer in the touchpoint and the customer's profile information.

Interact is closely integrated with IBM Campaign (see Chapter 6, "Leveraging Customer Data within IBM Campaign" on page 103) to define which offers are assigned to which customer. Because of this integration, marketers can use the same offers across all of your campaigns, along with all of IBM Campaign's offer management tools. Users can also integrate all the contact and response history across all of marketing campaigns. For example, email and direct mail contacts can influence offers presented to the user in real time.

The following sections describe IBM Interact components and how they work together.

7.3 IBM Interact: Architectural overview

Understanding the IBM Interact architecture can help clarify how Interact communicates with the customer-facing touchpoints, the runtime servers, and IBM Campaign. The Interact API leverages the runtime environment, and sometimes the testing runtime environment, to meet your performance requirements.

Interact consists of two major components: the design environment and the runtime environment. Marketers might also have optional testing runtime environments.

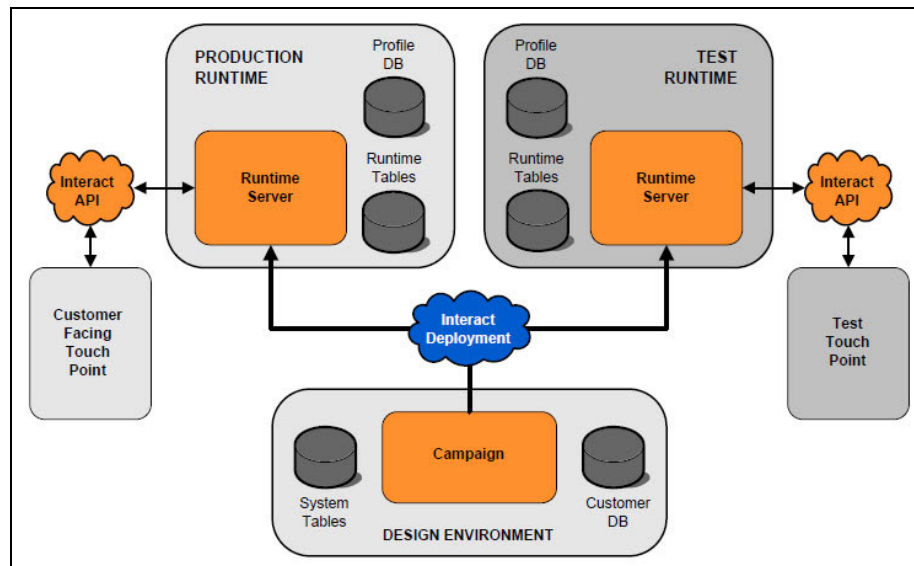Figure 7-1 shows the high-level architecture overview.

Figure 7-1   IBM Interact overview

The design environment is where logic is designed to determine how Interact makes its decisions. The design environment is installed with IBM Campaign and references the Marketing Platform and Campaign system tables, and also customer databases and other operational data sources.

After designing and configuring how Interact will respond to customer interactions, the design plan is deployed to either a testing runtime environment for testing or a production runtime environment for real-time customer interaction.

In production, the architecture might be more complex. For example, a runtime environment might have several runtime servers connected to a load balancer to meet failover and performance requirements.

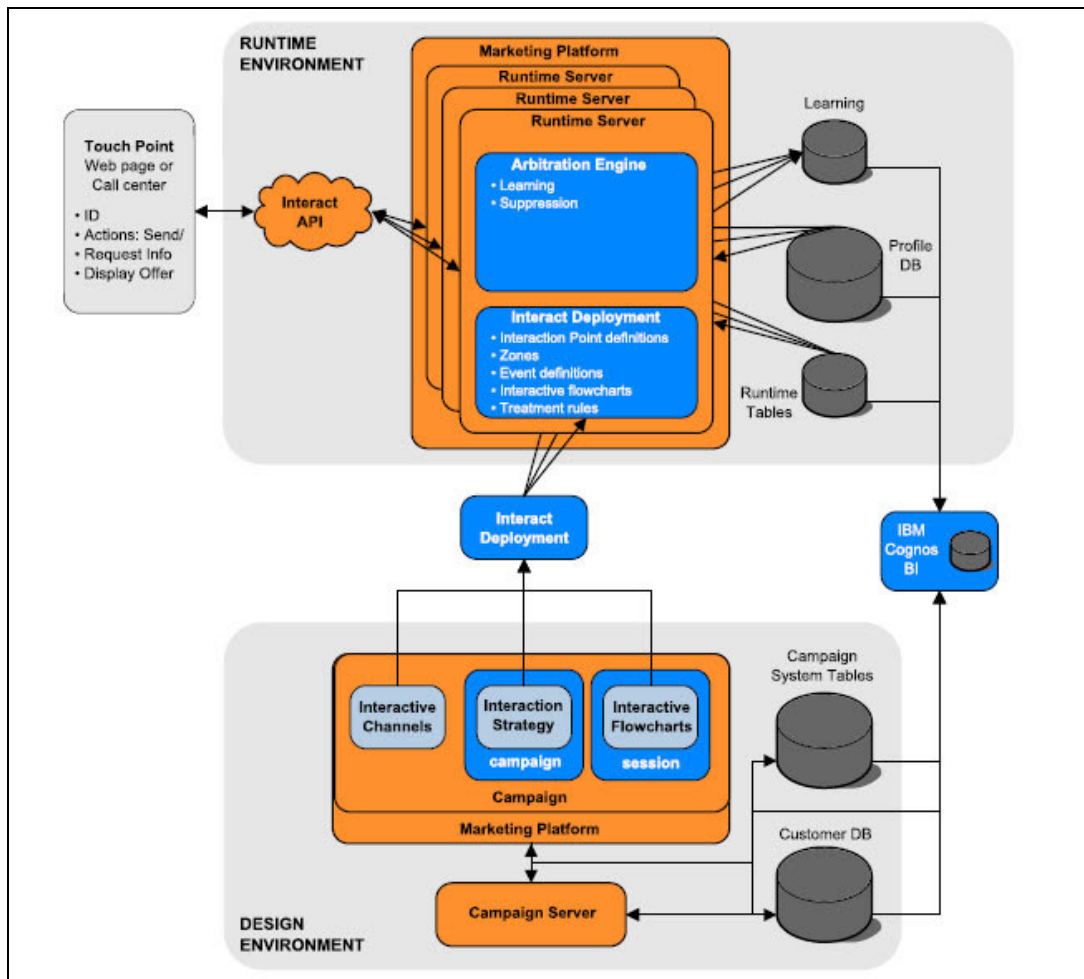Figure 7-2 shows the Interact environment in more detail.

Figure 7-2   Interact environment

In the design environment, users define how Interact responds within certain areas in the touchpoint by configuring interactive channels. Customers are then divided into segments by creating interactive flowcharts. The use of flowcharts borrows on the flowchart engine within IBM Campaign; therefore, IBM Campaign users will already be familiar with this paradigm. Within the interactive flowcharts, marketers can perform test runs to confirm that your customer data is segmenting correctly.

Next, users must define offers to align with the segmentation and assign the offers to segments within an interaction strategy. After users configure all of the

Interact components, they are ready to deploy the configuration to a test or production runtime environment.

IBM Interact deployments consist of the following items:

► Interact configuration data including interactive channels and interaction strategies.

► A subset of Campaign data including smart segments, offers, and interactive flowcharts.

► A customer data source to be used within Interact to make decisions in real time.

Because of the constraints of a real-time environment, response time is a critical factor in ensuring the user experience is acceptable. It is often not practical to mine a large customer database for decision data to be used in real time. As demonstrated in the preceding chapters, customers should design a subset of the data needed for real-time decision-making. This is usually a more flattened and less normalized view of key customer attributes including information such as demographics, areas of interest, loyalty levels, purchase history, and so forth.

In the test runtime environment, which mirrors the production runtime environment, customers validate the entirety of the Interact configuration, including the API integration with the touchpoint. During run time, a customer, or in the case of a test server, someone testing the system, takes actions in the touchpoint. These actions send events or requests for data to the runtime server by means of the Interact API. The runtime server then responds with results, such as presenting a set of offers (data) or re-segmenting a customer into a new segment (event). Users can continue modifying the Interact configuration in Campaign and redeploying it to the test runtime environment until they are satisfied with the behavior. Users then deploy the configuration to the production runtime environment.

The production runtime servers record statistical and historical data such as contact history and response history. The runtime contact response history is periodically written back into the IBM Campaign contact response tables in order to consolidate the information from multiple run-time environments and to make the information available to the entire cross-channel marketing platform. It also provides control over the size of the runtime contact response history data stores.

Contact Response data is used in reports to determine the effectiveness of the marketing offers driven by Interact and revise the associated decision logic as necessary. This data is also used by Campaign and other IBM cross-channel marketing products such as Contact Optimization, by integrating your real-time campaigns with traditional campaigns. For example, if a customer has accepted

an offer on the website, the offer acceptance data can be used in Campaign to ensure that either the same offer is not sent by mail, or the same offer is not used in a follow-up interaction. See Figure 7-3.
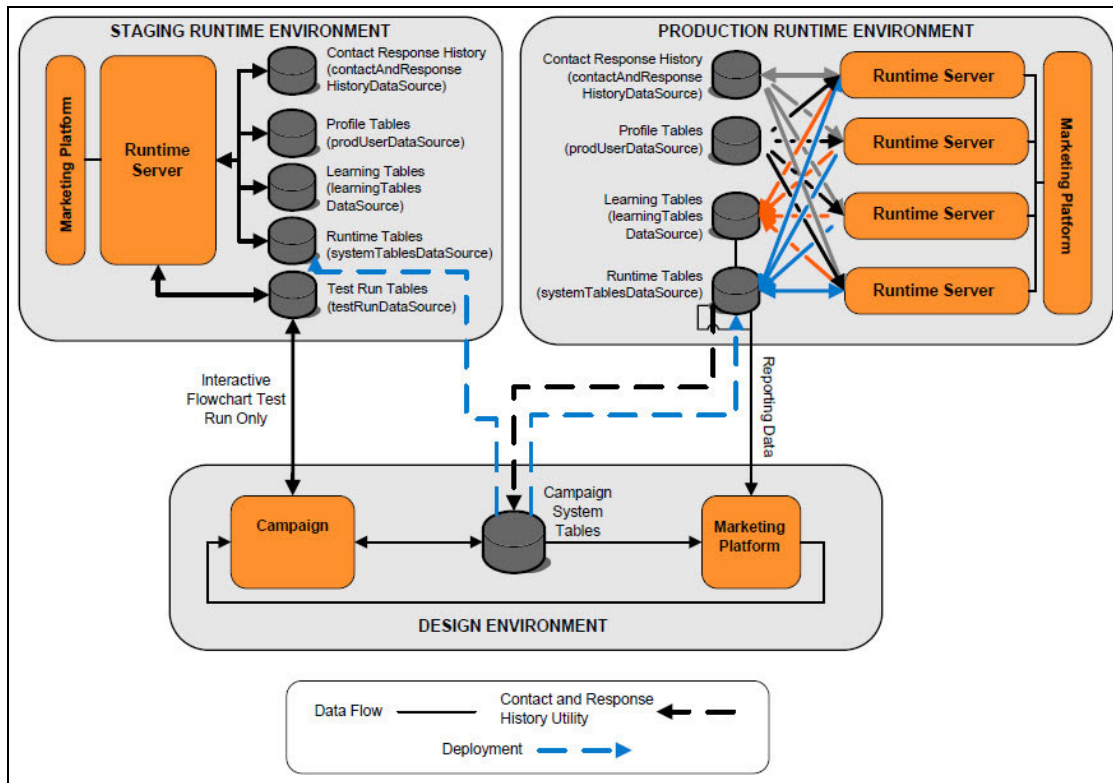

Figure 7-3   Contact response

7.4 IBM Interact: Process components

This section describes several key concepts you should understand before working with IBM Interact.

Interact integrates with IBM Campaign to define the offers assigned to customers. Integration between Interact and Campaign ensures that offers are consistent across multiple campaigns, along with the integration of contact and response history across all campaigns.

7.4.1 Design environment

The Interact design environment is used to configure various Interact components and deploy them to the runtime environment.

The design environment is where marketers develop the segmentation and offer logic invoked by the customer endpoint application. Marketers define events, interaction points, smart segments, and treatment rules. After configuring these components, they are deployed to the runtime environment and are immediately available to the associated applications.

IBM Campaign is a prerequisite for the installation of the Interact Design Environment. The design environment is a plug-in to IBM Campaign.

7.4.2 Interactive channels

An interactive channel is a representation in Campaign of a touchpoint where the method of the interface is an interactive dialog. This software representation is used to coordinate all of the objects, data, and server resources that are involved in interactive marketing.

Interactive channels are used to coordinate all objects, data, and server resources that involved in interactive marketing. It defines interaction points and events.

Interactive channels also contain production runtime and staging server assignments. Marketers can create several interactive channels to organize events and interaction points if there is only one set of production runtime and staging servers, or divide events and interaction points by customer-facing system.

Reporting on an interactive channel is accessible from the Analysis tab of the interactive channel form.

7.4.3 Interactive flowcharts

An *interactive flowchart* is related to, but slightly different from, a Campaign batch flowchart. Interactive flowcharts contain a subset of the batch flowchart processes, and interactive flowchart-specific processes.
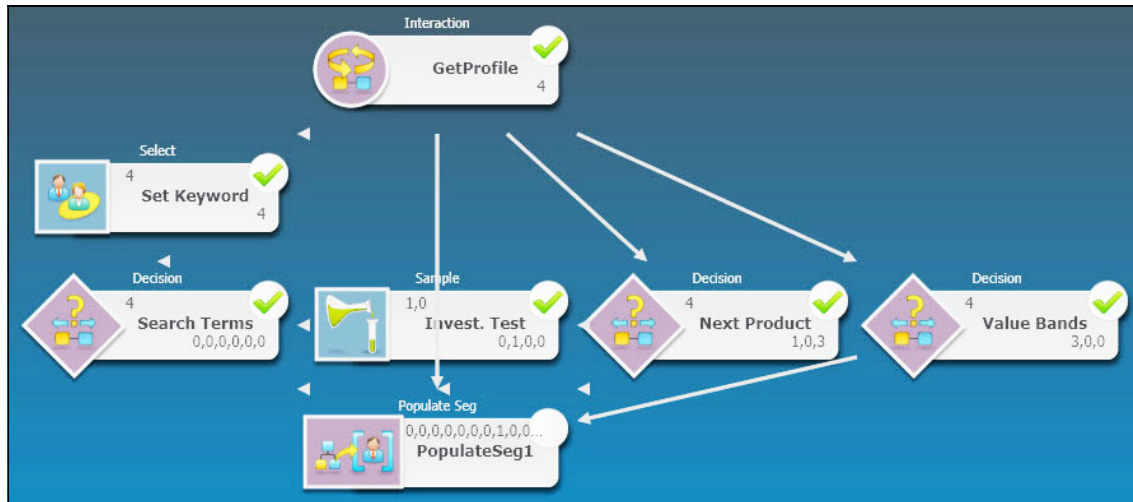

Figure 7-4   Interactive flowchart example

Interactive flowcharts perform the same major function as batch flowcharts: dividing customers into groups known as *segments*. However, for interactive flowcharts the groups are smart segments, which are populated at the time the flowchart runs. Interact uses these interactive flowcharts to assign a profile to a segment when a behavioral event or system event indicates that a visitor re-segmentation is needed.

Use interactive flowcharts to divide your customers into segments and then assign offers to those segments.

7.4.4 Interaction points

An *interaction point* is a place in the touchpoint channel where users want to present an offer. Interaction points are typically referred to as "real estate" or "personalization zones" within a web page or application window or screen. They are most often the location within a user interface where the interactive marketing messages are to be rendered or displayed. See Figure 7-5.
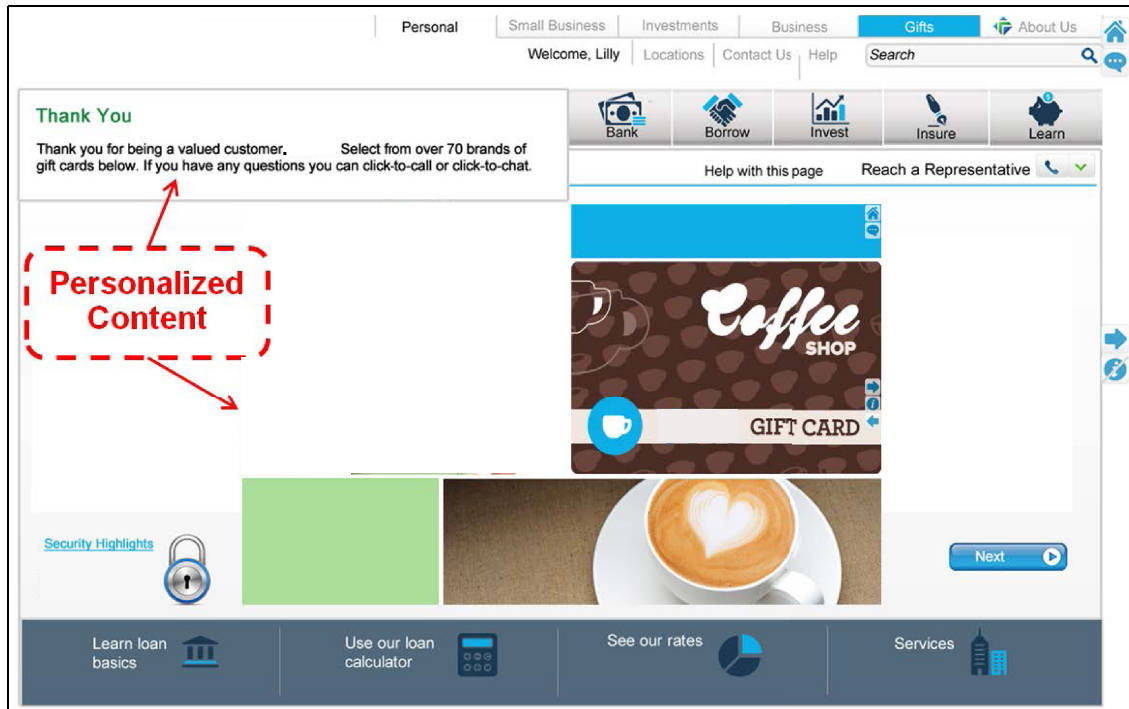

Figure 7-5   Personalization example

Interaction points contain default content in cases for when specific offer or eligible content might not be available for the user. Interaction points can be organized into zones to ease administration (examples include main banner or header, right sidebar, and footer).

7.4.5 Events

An *event* is an action that is taken by a visitor that then triggers an action in the runtime environment. Examples of an event might be placing a visitor into a segment, presenting an offer, or logging data. See Figure 7-6.

Figure 7-6   Events example

Events are first defined in an interactive channel and then triggered by a call to the Interact API by using the postEvent method. An event can lead to the execution of one or more actions that are defined in the Interact design environment. Examples of events might be contact, accept, reject, and trigger re-segmentation.

Triggering re-segmentation causes the associated decision flowchart to be rerun based on information that is provided through the API in order to change the flow of the decision logic. For example, Interact can be called by providing name and value pairs that indicate a department or product line in order to request different offers, based on where a user is currently browsing in an online store.

> **Note:** When designing interactions, remember that unless users specify a specific flowchart within the API call, a re-segmentation action will rerun *all* interactive flowcharts associated with this interactive channel for the current audience level again. Any request for offers wait until all flowcharts have completed. Excessive re-segmentation within a single visit can negatively affect the performance of the touchpoint and the customer's experience.

Customers often find their way into new segments after significant new data is added to the runtime session object, such as new data from requests from the Interact API (such as changing the audience) or customer actions (such as adding new items to a wish-list or shopping cart).

The following list describes examples of events and actions:

► Log Offer Contact: The runtime environment flags the recommended offers for the database service to log the offers to contact history.

   For web integrations, log the offer contact in the same call where users request offers to minimize the number of requests between the touchpoint and the runtime server.

   If the touchpoint does not return the treatment codes for the offers, which Interact presented to the visitor, the runtime environment logs the last list of recommended offers.

► Log Offer Acceptance or Rejection: The runtime environment flags the selected offer for the database service to log to response history.

► Trigger User Expression: An expression action is an action that you can define by using Interact macros, including functions, variables, and operators, including EXTERNALCALLOUT. You assign the return value of the expression to any profile attribute. This eases the integration of external functions that might not be part of the ready-to-use product.

   When clicking the edit icon next to Trigger User Expression, the standard User Expression editing dialog opens, and you can use this dialog to specify the audience level, optional field name to which to assign the results, and the definition of the expression itself.

► Trigger Events: Use this action to enter an event name that you want to be triggered by this action. If you enter an event that is already defined, that event is triggered when this action is run. If the event name enter does not exist, this action causes the creation of that event with the specified action.

Users can also use events to trigger actions that are defined by the postEvent method, including logging data to a table, including data to learning, or triggering individual flowcharts.

Events can be organized into categories for your convenience in the design environment. Categories have no functional purpose in the runtime environment.

7.4.6 Profiles

A *profile* is a customer data source that is used by the runtime environment. This data can be a subset of the customer data available in your customer database, data that is collected in real time, or a combination of the two. In the following example, IBM Interact will be using data from the Marketing Data Mart, as described in Chapter 3, "Configuring MDM with marketing-relevant attributes" on page 27.

The customer data is used by the flowchart logic to make decisions. This data may be used for the following reasons:

► To assign a customer to one or more smart segments in real-time interaction scenarios

   A profile data source is required for each audience level by which segmentation is defined. For example, if you are segmenting by location, you might include only the customer's postal code from all the address information you have.

► To personalize offers

► As attributes to track for learning

   For example, users will configure Interact to monitor the marital status of a visitor and how many visitors of each status accept a specific offer. The runtime environment can then use that information to refine offer selection.

The profile data source is read-only within the runtime environment. Data collected as a result of an interaction should be written as a transaction and processed by the central customer data mart.

7.4.7 Runtime environment

The runtime environment connects to touchpoints and responds to real-time interaction requests. The runtime environment can consist of one or many runtime servers that are connected to a touchpoint.

The runtime environment uses the information that is deployed from the design environment in combination with the Interact API to retrieve offer data to be rendered within the calling application or touchpoint.

7.4.8 Runtime sessions

A runtime session is created on the runtime server the first time an Interact API request is made. The session persists until it is explicitly ended or times-out based on the Interact server configuration.

Customer profile data and flowcharts are loaded into memory for the duration of the session, speeding execution and response time.

You create a runtime session when you use the startSession call. The call returns a sessionID that persists for the duration of the session.

7.4.9 Smart segments

A *smart segment* is similar to a strategic segment (IBM Campaign) in that it is a group of customers with defined traits. Instead of a list of IDs, however, a smart segment is the definition of what IDs are allowed in the list.

For example, a smart segment might be as follows:

"All customers living in Colorado with an account balance greater than $10,000 who have applied for a car loan in the last 6 months."

These definitions are represented by interactive flowcharts. Smart segments are available only in Interact.

7.4.10 Touchpoints

A *touchpoint* is an application or place where you can call IBM Interact to provide an offer or message to a customer. A touchpoint is generally a channel where the customer initiates the contact (an "inbound" interaction) or where you contact the customer (an "outbound" interaction). Common examples are websites, call center applications, and interactive voice response (IVR) systems.

Using the Interact API, you can integrate Interact with your touchpoints to present offers to customers based on their action in the touchpoint. Touchpoints are also sometimes called client-facing systems (CFS).

7.4.11 Treatment rules

Treatment rules assign an offer to a smart segment. These assignments are further constrained by the custom-defined zone that you associate with the offer in the treatment rule.

Each treatment rule also has a marketing score. If a customer is assigned to more than one segment, and therefore more than one offer is applicable, the marketing scores help define offer priority. Which offers the runtime environment suggests can be influenced by a learning module, an offer suppression list (blacklist), and global and individual offer assignments (whitelist).

7.4.12 Interact API

Use the Interact API to integrate Interact with your touchpoints. The Interact API can be invoked using serialized Java, SOAP, or REST to integrate Interact with your touchpoints.

7.4.13 Zones

Interaction points are organized into zones. You can limit a treatment rule to apply to only a certain zone.

If you create a zone that contains all of your "welcome" content, and another zone for "attrition risk" content, you can present a different set of offers to the same segment based on where the customer is in your touchpoint.

7.5 IBM Interact: User workflow

When you configure Interact, you use a workflow to outline the process from an idea to deployment in detail. Configuring Interact is a multi-step, multi-person, iterative process. The process from an idea to deployment can be divided into three major components:

► Design
► Configuration
► Testing

7.5.1 Design

During the design phase, you brainstorm about what kinds of interactive marketing strategies you want to use. After you have an idea about what you want to happen in your touchpoint, determine how to implement that with Interact.

This brainstorming is a cooperative effort between the person managing the touchpoint and the person designing the marketing plan. Using business goals and target metrics, they can create a list of interaction points and zones, and a rough list of segmentation and suppression strategies. These discussions should also include the data that is required to do the segmentation and, in our use case, include a real-time call to the Marketing Repository for the most current customer data.

7.5.2 Configuration and development

During the configuration phase, the touchpoint administrator and an Interact user implement the design. The Interact user defines offer-to-segment assignments and configures the interactive channel within the design environment while the touchpoint administrator configures the touchpoint to work with the runtime server using the Interact API.

Your data administrator needs to configure and create the data tables required for both testing and production. Depending on the application, the touchpoint administrator might need to be an IT or web developer, or programmer.

7.5.3 Testing

After all development in the design environment is complete, you mark the various components for deployment to test the runtime environment. The Interact administrator deploys the configuration to the appropriate servers and the testing can begin.

All team members who are involved with designing the Interact implementation review the results. The purpose is to confirm that the configuration is performing as designed and the performance of the runtime environment is within tolerable limits for response time and throughput.

Users might need to make some changes and more testing might need to be done. When everyone is satisfied with the results, the manager can mark the configuration for deployment to the production environment.

At this time, the team will likely retest in production to ensure that the configuration has no negative effects on the systems or customer experience. If a problem is identified, the deployment can be instantly rolled back from production.

Figure 7-7 on page 144 shows a sample design workflow. Although this diagram shows a linear progression, in practice, many people can be working on different components at the same time. This is also an iterative process. For example, to configure the touchpoint to work with Interact using the Interact API, you must reference events created in the interactive channel. As the touchpoint administrator configures the touchpoint in the runtime environment, the administrator might realize that more events are needed. The Interact user then must create these events in the design environment.
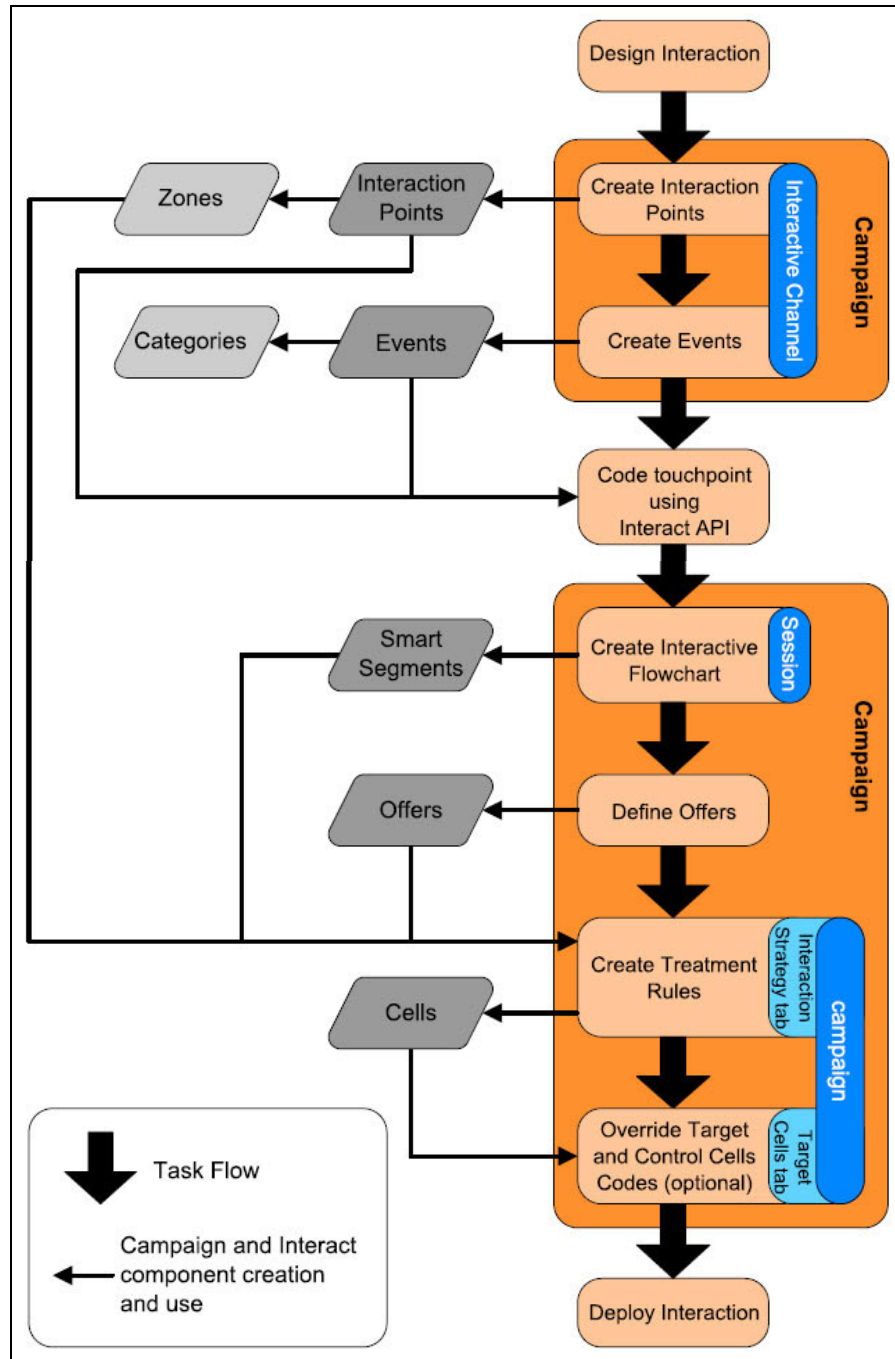
Figure 7-7 Sample design workflow

Tasks are as follows:

1. The first task is to design the interaction.

2. Next, create interaction points, zones, events, and categories in the interactive channel. The touchpoint administrator uses the names of the interaction points and events with the Interact API to configure the touchpoint.

3. Continue to configure the interaction, creating interactive flowcharts, and defining offers.

4. After you create all your zones, offers, and segments, you can create treatment rules on the Interaction Strategy tab of a campaign. This is where you assign offers to segments per zone. You might spend some time to assign target and control cells on the Target Cells tab.

The interaction is now complete and can be deployed to a staging server for testing.

7.6 IBM Interact: Mapping tables

Mapping tables is the process of making external customer or system tables accessible within IBM cross-channel marketing.

A table mapping generates metadata that the system uses to define a base, dimension, or general table. It contains information about the data source, the table name and location, the table fields, audience levels, and data. Table mappings can be stored for reuse in table catalogs.

The contents of the profile database depend entirely on the data you need for configuring your interactive flowcharts and Interact API. Interact requires or recommends that each database contain certain tables or data.

The profile database must contain all tables mapped in the interactive channel. These tables must contain all the data required for running your interactive flowcharts in production. These tables should be flattened, streamlined, and properly indexed. Because there is a performance cost to access dimensional data, use a denormalized schema when possible. At a minimum, index the profile table on the audience level ID fields. If other fields are retrieved from dimensional tables, these should be indexed appropriately to reduce database fetch time. The audience IDs for the profile tables must match the audience IDs defined in IBM Campaign.

Map the profile table for an interactive channel

Use the Summary tab on an interactive channel to map the profile table for the interactive channel. You must map a profile table before you can map any dimension tables. Use the following steps:

1. Click **unmapped** for the audience level you want to map under Mapped Profile table on the interactive channel Summary tab. The Map Profile tables for the Audience Level wizard opens.

2. Complete the Map Profile tables for Audience Level wizard.

When the Map profile tables for Audience Level wizard validates the table mapping, it references the data source that is defined in the **Campaign** → **partitions** → **partitionN** → **Interact** → **flowchart** → **data source** property. All interactive channels must reference the same data source. This data source is for test runs only.

To edit the mapping of a profile table, click the name of the Mapped Profile Table and complete the Map Profile tables for Audience Level wizard. This is depicted in Figure 7-8.
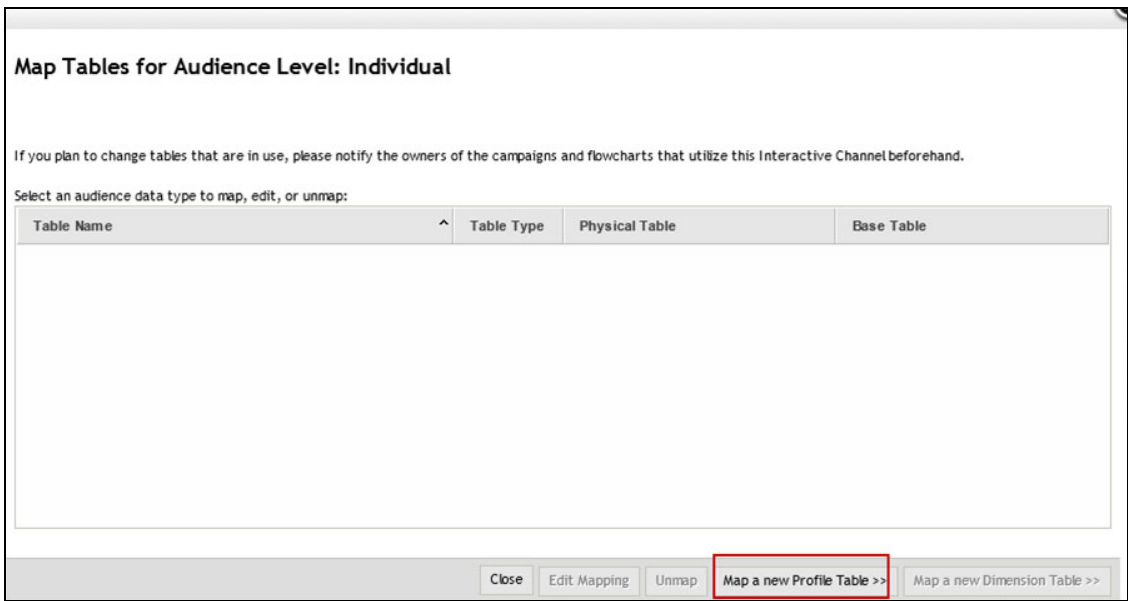
Figure 7-8   Map a new table profile

To map a new profile table into Interact, use the wizard to complete these steps:

1. Click **Map a new Profile Table**.

2. Navigate to and select **dbo.mdm4mktg** (Figure 7-9), and then click **Validate Mapping**.

Figure 7-9   Validate Mapping

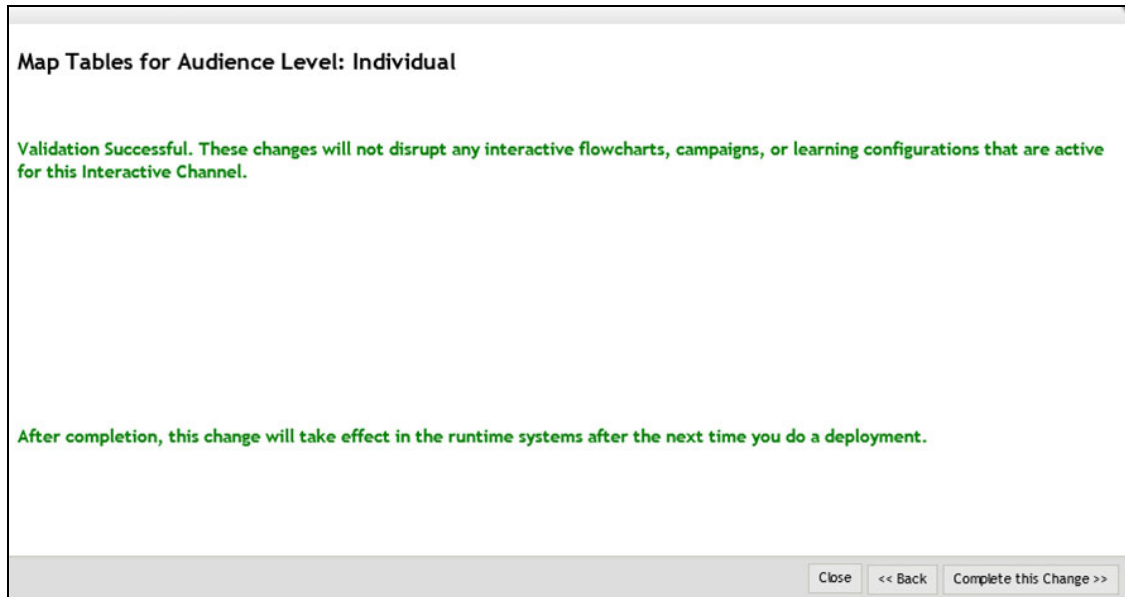3. Click **Complete this Change**. See Figure 7-10.

Figure 7-10   Complete the change

Map dimension tables for an interactive channel

Use the Summary tab of an interactive channel to map dimension tables for the interactive channel. You must map the profile tables for the interactive channel before you can map the dimension tables.

The Map profile tables for Audience Level wizard references the data source that is defined in the **Campaign → partitions → partitionN → Interact → flowchart → data source** property

All interactive channels must reference the same data source. Complete the following steps:

1. Click the name of a profile table under Mapped Profile table on the interactive channel Summary tab. The Map Profile tables for Audience Level wizard opens.

2. Click **Map a new Dimension Table**. See Figure 7-11 on page 149.

3. Complete the Map Profile tables for Audience Level wizard.

Figure 7-11   Map new dimension table

4. Navigate to and select **dbo.mdm4mktg_contact** (Figure 7-12), and then click **Map Dimension to Base**.

Figure 7-12   Mapping dimensions to the base

5. Click **Complete this Change**. See Figure 7-13.
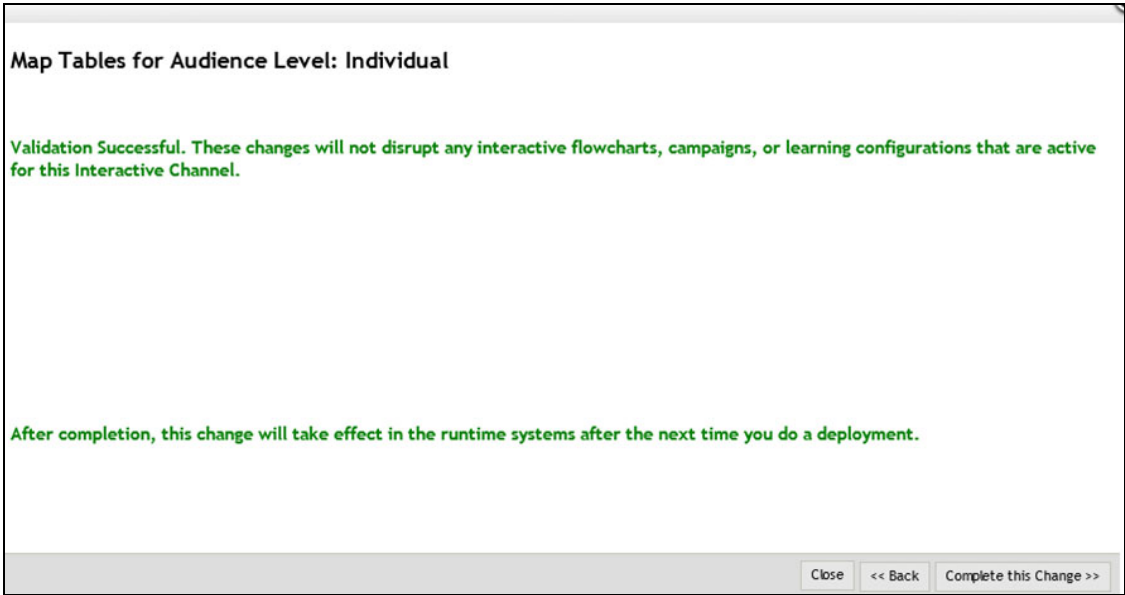

Figure 7-13   Complete the change

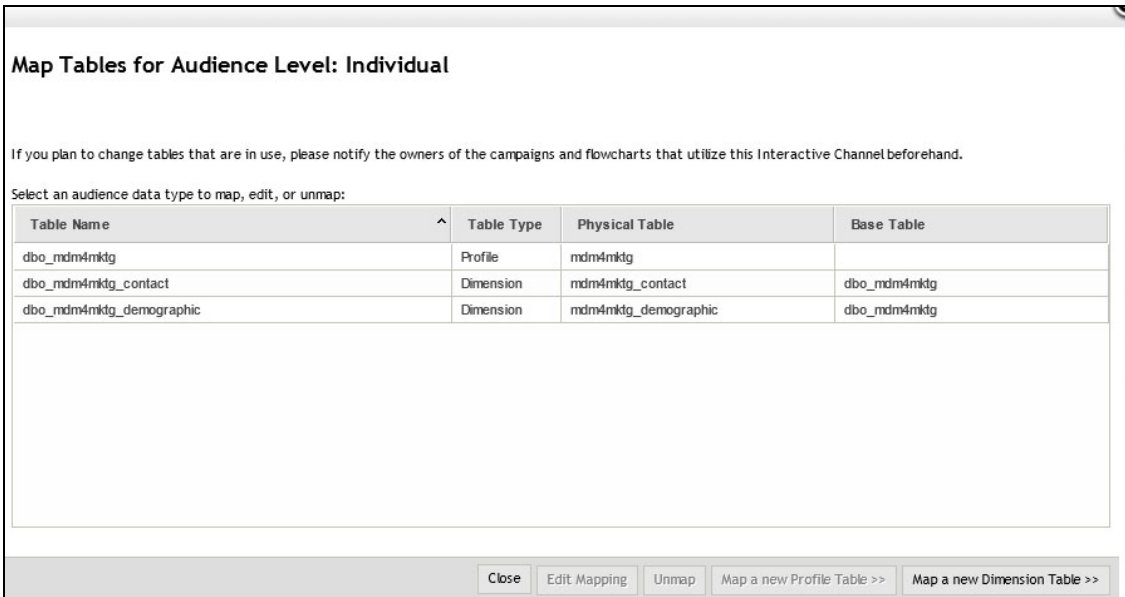6. Mapping is complete (Figure 7-14). Click **Close**.

Figure 7-14   Close the process

7.7 IBM Interact: Profile Data Services

The Profile Data Services API is contained in the iInteractProfileDataService interface. This interface allows you to import hierarchical data into an Interact session through one or more external data sources (such as a flat file, web service, and so on) at the time the Interact session starts or the audience ID of an Interact session changes.

To develop hierarchical data import using the Profile Data Services API, you must write a Java class that pulls information from any data source and maps it to an ISessionDataRootNode object. Then, you refer to that mapped data using the EXTERNALCALLOUT macro in a Select process of an interactive flowchart.

You must compile your implementation against the following file, which is located in the `lib` directory of your IBM Interact runtime environment installation:

`Interact_externalcallout.jar`

For a complete set of Javadoc documentation for using this interface, view the files in `Interact_home/docs/externalCalloutJavaDoc` with any web browser.

For a sample implementation of how to use the Profile Data Service, including commented descriptions of how the example was implemented, see this file:

`Interact_home/samples/externalcallout/XMLProfileDataService.java`

7.7.1 Add a data source for use with Profile Data Services

The EXTERNALCALLOUT macro recognizes a data source for Profile Data Services hierarchical data import only if you have defined the appropriate configuration properties.

In Marketing Platform for the runtime environment, add or define the following configuration properties in the **Interact** → **profileAudience Levels** → **[AudienceLevelName]** → **Profile Data Services** category. See Table 7-1 on page 152.

Table 7-1   Configuration properties

Property	Setting
New category Name category	Indicates the name of the data source you are defining. The name you enter here must be unique among the data sources for the same audience level.
enabled	Indicates whether the data source is enabled for the audience level in which it is defined.
className	This is the fully qualified name of the data source class that implements iInteractProfileDataService interface.
classPath	This is the class path to your Profile Data Services class files. If you omit it, the class path of the containing application server is used by default.
priority category	This is the priority of this data source within this audience level. It must be a unique value among all data sources for each audience level. (That is, if a priority is set to 100 for a data source, no other data source within the audience level may have a priority of 100.)

7.8 Conclusion

This chapter introduces the process that marketers can follow in using the most accurate customer data to produce a real-time inbound campaign that is designed to reduce attrition risk for a specific group of customers.

8

Where do you go from here

So you have created an information supply chain that takes your trusted customer information and feeds a data warehouse, which in turn populates a marketing repository for your outbound marketing using IBM Campaign.

You have exposed the customer profile from MDM as a service, so your inbound marketing can get real-time information.

The previous chapters describe how to establish a foundation for trusted customer information for inbound and outbound marketing. The trusted customer information is based on IBM Information Management capabilities including master data management, data integration and governance, and industry models. The inbound and outbound marketing is based on the IBM Enterprise Marketing Management platform.

This chapter describes several ways in which you can build upon the EMM platform to keep improving and optimizing the way you manage information, analyze that information for decision-making, and interact with customers.

8.1 Managing your information

If you are to deliver individualized, timely, relevant marketing that you can be confident in, then you need to know who you are sending offers to, how you know that the offer is appropriate, and whether you can easily determine the basis on which the offer was made. How can you be sure of this, when offers are made in real time, and using data from many sources that are constantly changing? How you gather, manage, store, and move and govern information becomes more important to be able to provide the right information, at the right time, to make the right offer to the right person. This is true for the offers you present to your customers, but also for all the other internal and external roles touched by your enterprise. This section describes some ways you can enhance your information supply chain to meet these needs.

8.1.1 Big Data

You are in a new age, where there are many new types of data from the Internet, blogs, tweets, other social media, reviews, and comments. This data is often large in volume, full of opinion, collaborative and unstructured, and accessible everywhere on mobile devices. Influences on people's opinions occur quickly, so a snapshot of this information might have a small time frame in which it is relevant. It is in this world that buying decisions are made. Your marketing efforts need to fit in with where your customers and prospects are and what they are doing. Big data can be characterized as having extremes of some or all of the following four dimensions: volume, velocity, variety, and veracity.

See the following image for more information:

http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

One way to use this data is to capture it in stores such as Apache Hadoop (and IBM Hadoop Systems), then make sense of it by performing text analytics, for example by using the IBM SPSS® Analytic Server. Another way is to use InfoSphere Streams to look at the large flow of data, spot the interesting data, and perform analytics on it.

For more information. see the following sources:

▶ Big Data at the Speed of Business website:

http://www.ibm.com/software/data/bigdata/

▶ *Smarter Analytics: Information Architecture for a New Era of Computing*, REDP-5012

8.1.2 Multi-domain MDM

In this book, you considered Party domain in MDM, which contains people and organizations. MDM can also manage other domains of data. A fruitful next step might be to set up a product domain, where you centrally store your product data in the MDM operational hub. You might also consider storing your accounts in the account domain. Imagine the power of having a central view of customers and everything they have bought and the terms and conditions associated with the account in an operation hub available in real time. Not only do you have these high-value business information sources stored centrally, but also the relationships between them, allowing you to manage lifecycles of each of these types of master data coherently and efficiently.

As you build your MDM capabilities, the master data becomes of more value across your enterprise because it has a single view of the trusted high business value information in an operational hub.

For more information, see the following sources:

► IBM Master Data Management:

   http://www.ibm.com/software/data/master-data-management/

► *Enterprise Master Data Management: An SOA Approach to Managing Core Information* by Allen Dreibelbis, Eberhard Hechler, Ivan Milman, Martin Oberhofer, Paul Van Run, Dan Wolfson, published by IBM Press

8.1.3 Information governance

As you embrace greater volumes of data, from an increasing variety of sources and store more in your master data repository, it becomes important to understand the data lineage of the data in your enterprise and also the governance policies with which you must comply. Master data management allows you to control the lifetime and quality of your master data; when information is "well groomed" it has policies and processes around it that provide governance. A strong governance story allows you to be able to define, implement, and enforce policies in line with legal and ethical considerations. For example, when you store data about people, you must be careful about what data you store, what permissions were sought and obtained, who has access to it and masking private data, so that data is not inadvertently exposed.

For more information about IBM Information Integration and Governance, see the following website:

http://www.ibm.com/software/data/information-integration-governance/info-governance.html

8.1.4 Information virtualization and data lakes

With the move toward cloud computing, having information resources exposed in a virtual way allows data scientists and marketers to get their data the way they want it, without needing to know where it is. Data lakes manage and govern repositories of information for analytics use cases.

8.2 IBM Smarter Commerce

IBM Smarter Commerce™ is a set of rich capabilities geared around the customer to enhance buying, selling, marketing, and service to create an integrated value chain. This book describes IBM Campaign and IBM Interact, two products that are part of Enterprise Marketing Management (EMM). Figure 8-1 shows how these products fit in with EMM and Smarter Commerce.
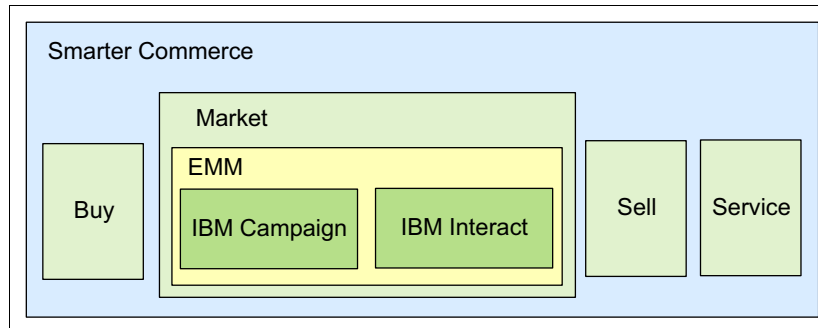

Figure 8-1   EMM and Smarter Commerce

To further enhance IBM Interact and gain insight, see IBM Digital Analytics:

http://www-03.ibm.com/software/products/en/digital-analytics/

With smarter commerce, things start with the customer. Having a single trusted view of a customer across your enterprise becomes a foundation on which to build your smarter commerce solution. Whether you are tying together supply chains, tracking products with Product and Accounts domains, or just making sure you know who you are marketing to and whether there was a subsequent sale with the Party domain, MDM has a key role to play.

For more details about Smarter Commerce, see the following website:

http://www.ibm.com/smarterplanet/uk/en/smarter_commerce/overview/

8.3 Embracing analytics

Analytics are increasingly becoming the differentiator between you and your competitors. Successful companies are using analytics and data-driven decisions. Folding analytic approaches into your EMM gives you the next level of insight into your customers and prospects.

For more information about IBM business analytics capabilities, see the following website:

http://www-03.ibm.com/software/products/en/category/SWQ00

Analytics can be categorized as descriptive, predictive, prescriptive, and cognitive.

8.3.1 Descriptive analytics

Using business intelligence (BI), you can get real insights on your customers and their buying behaviors. These insights can be used as a basis for marketing campaigns.

8.3.2 Predictive analytics

Basing the marketing decisions on the way people have behaved in the past, you can use data mining techniques to understand how they might behave in the future. Given enough data (and there are masses of data in the world of social media), a data scientist can train an analytic model that can then predict values for new records and give a probability of the prediction being correct. For example, if the enterprise is concerned with customers leaving the company, and they have data about all their customers and what those customers have done, including which ones left, then a data scientist can use this information to train an analytics model to determine who is likely to churn.

When you have predictive analytics, you want the system to be able to act on that information when making a decision. For example, if you have churn information, you can use it in your marketing campaigns to take actions to reduce the chances of a customer leaving. Using this data-driven approach, you are much more likely to be able to give individualized offers and actions to encourage the desired action. In this way you are moving closer to the segment of one.

The next best action (described in 8.4, "Next best action" on page 159) uses predictive analytics.

8.3.3 Prescriptive analytics

The developerWorks blog defines prescriptive analytics as "A set of mathematical techniques that computationally determine a set of high-value alternative actions or decisions given a complex set of objectives, requirements, and constraints, with the goal of improving business performance."

See the following web page:

http://bit.ly/1hSGZXb

8.3.4 Cognitive analytics

Cognitive systems, such as IBM Watson™, have the ability to parse natural language in real-time and make decisions. So, after Watson is trained in your domain, you can ask it questions in speech and get intelligent answers!

For more information see the following sources:

► IBM Watson:

  http://www-05.ibm.com/uk/watson/

► *Smarter Analytics: Taking the Journey to IBM Cognitive Systems*, REDP-5043

8.4 Next best action

The next best action is a real-time analytics solution, containing marketing and the customer master. The pattern includes a system of record exposed in operational hubs, for use by real-time analytics. It shows how various capabilities can work together, including big data, information gathering, marketing, decisions, and reporting, and how to coordinate the flow of information between them. Using this pattern, you can decide on a road map for how you want to move toward real-time analytics based on trusted information.

The next best action pattern in Figure 8-2 gives you an overview of the software capabilities that you might need and how they can work together.
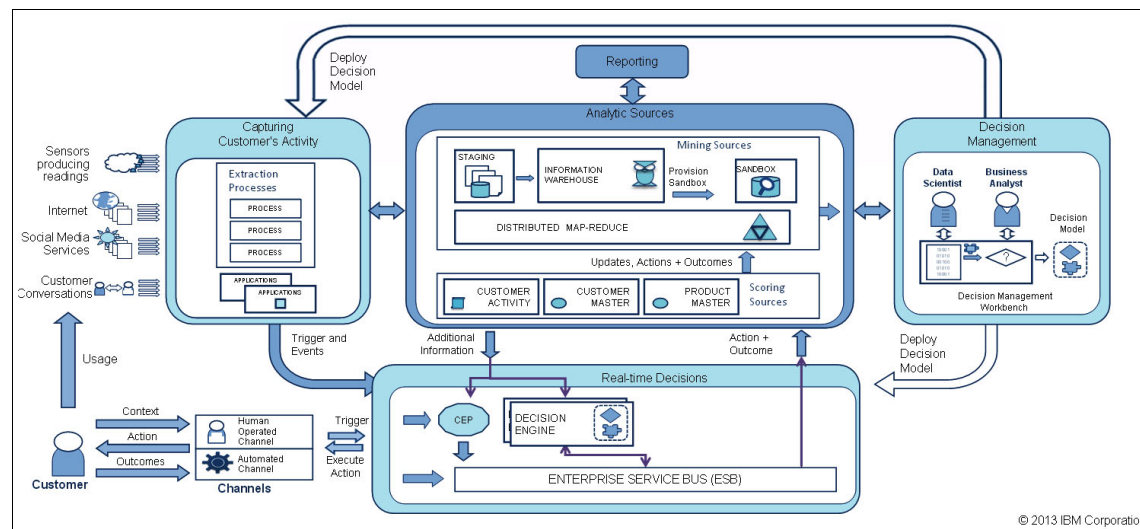
Figure 8-2   Next best action

For more details, see *Smarter Analytics: Driving Customer Interactions with the IBM Next Best Action Solution*, REDP-4888-01.

If you want to capitalize on the power of analytics in your business, you do not need to implement all of what is shown in Figure 8-2 before you can gain value. You can gradually adopt good practices for creating trusted data and building in analytic insight. Here are some considerations as you proceed:

► You must be clear about your marketing and analytic goals. What is your business problem and how will you know when you have addressed it?

► You must understand the information integration required for your use cases to be able drive effective analytics.

- You can gain much insight by running analytics on a historical store. This insight can inform your marketing and decisions. See the *Cross Industry Standard Process for Data Mining (CRISP-DM)* guide:

  `ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf`

- Scoring analytics looks for two types of insight:
  - New facts about the customer
  - Potential triggers for marketing actions

- When insight is generated from the text analytics, it may be new facts or triggers for action.

- When a decision is made, the inputs and outcome must be preserved to tune the analytic model.

8.5 Summary

This is an exciting time in the history of marketing and analytics. You are now better enabled to better understand where and how marketing works with big data, traditional data, social media, cloud technology and mobile. As you embrace and use these new technologies, you will need the backing of trusted governed data and information integration.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

► *Smarter Analytics: Driving Customer Interactions with the IBM Next Best Action Solution,* REDP-4888

► *Smarter Analytics: Information Architecture for a New Era of Computing*, REDP-5012

► *Smarter Analytics: Taking the Journey to IBM Cognitive Systems*, REDP-5043-00

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

► Mandy Chessell and Harald Smith, *Patterns of Information Management*, IBM Press; 1 edition (16 May 2013), ISBN-10: 0133155501, ISBN-13: 978-0133155501

► Sunil Soares, *IBM InfoSphere: A Platform for Big Data Governance and Process Data Governance*, MC PRESS (31 Mar 2013), ISBN-10: 1583473823, ISBN-13: 978-1583473825

► Allen Dreibelbis, Eberhard Hechler, Ivan Milman, Martin Oberhofer, Paul Van Run, and Dan Wolfson, *Enterprise Master Data Management: An SOA Approach to Managing Core Information*, IBM Press; 1 edition (5 June 2008), ISBN-10: 0132366258, ISBN-13: 978-0132366250

Online resources

These websites are also relevant as further information sources:

► Enhancing IBM Enterprise Market Management solutions with master data A trusted base for EMM

   http://www.ibm.com/developerworks/data/library/techarticle/dm-1303enhance/index.html?ca=dat-/

► IBM InfoSphere Platform

   http://www.ibm.com/software/data/infosphere/

► IBM InfoSphere Master Data Management

   http://www.ibm.com/software/data/master-data-management/

► IBM InfoSphere Streams

   http://www-03.ibm.com/software/products/en/infosphere-streams/

► IBM Master Data Management v11 Information Center

   http://pic.dhe.ibm.com/infocenter/mdm/v11r0/index.jsp

► IBM MDM Developers community

   http://ibm.co/mdmgroup

► OSGi Alliance

   http://www.osgi.org/Main/HomePage/

► MDM Workbench

   http://ibm.co/PMjK9U

► IBM enterprise marketing management (EMM) Information

   http://www-03.ibm.com/software/products/en/category/SWX00/

► IBM Campaign

   http://www-03.ibm.com/software/products/en/campaign-management/

► IBM Interact

   http://www-03.ibm.com/software/products/en/real-time-inbound-marketing/

► IBM Digital Analytics

   http://www-03.ibm.com/software/products/en/digital-analytics/

► IBM Big Data

   http://www.ibm.com/software/data/bigdata/

- ► The FOUR V's of Big Data

  http://www.ibmbigdatahub.com/sites/default/files/infographic_file/4-Vs-of-big-data.jpg

- ► IBM business analytics

  http://www-03.ibm.com/software/products/en/category/SWQOO/

- ► IBM developerWorks blog

  http://ibm.co/1hxnaVn

- ► IBM SPSS predictive analytics

  http://www.ibm.com/software/uk/analytics/spss/

- ► IBM Watson

  http://www-05.ibm.com/uk/watson/

- ► IBM Cross Industry Standard Process for Data Mining (CRISP-DM)

  ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

IBM

Redbooks

Enhance Inbound and Outbound Marketing with a Trusted Single View of the Customer

®

Enhance Inbound and Outbound Marketing with a Trusted Single View of the Customer

Redbooks ®

Build on an IBM Enterprise Marketing Management Platform

Include business and IT functionality for marketing

Use an IBM industry standard data model

IBM Campaign and IBM Interact are critical components in an Enterprise Marketing Management (EMM) platform. They are the foundation for optimizing your marketing campaign effectiveness, marketing operations, and multi-channel marketing execution. However, the effectiveness of the marketing campaigns is highly dependent on the quality, accuracy, and completeness of the underlying customer information used by the EMM platform. IBM InfoSphere Master Data Management (MDM) is a trusted source of that complete, accurate, customer information. Using your master data as the basis for running marketing campaigns provides the best information available for the best possible return-on-investment for your marketing operations.

This IBM Redbooks publication describes how master data about customers is extracted from an MDM hub and delivered through an "information supply chain" to your marketing data repository. This information supply chain includes capabilities such as data integration, metadata management, industry data models, and workload-optimized analytics appliance. The intent of this book is to give marketing organizations (both the business and IT functions for marketing) a blueprint for how to architect your EMM solution in a way that best takes advantage of your trusted master data.