



DAMA MICHIGAN Bits & Bytes



Michigan Chapter of DAMA International

Winter 2002

Board of Directors

President

JoAnne Gierman
MI Dept. of Transportation

Administrative Vice President

Rajiv R. Das
Accenture

Vice President - Marketing

Bill R. Hepburn
TRW

Vice President - Membership

Carl Guse
Universal Solutions, Inc.

Vice President - Programs

Michelle Rymer - EDS
Yvette Biggs - GM

Treasurer

Sid Stoffer
EDS

Vice President

Online Communications

Jerry Stembridge
EDS

Secretary

Nancy Dwyer
EDS

Vice President -Logistics

John A. Mieczkowski
EDS

Corporate Sponsors



Dr. Vernon Hoffner to address DAMA Michigan Chapter January 8

January 8, 2002

5:30 p.m.—8:30 p.m.
Lansing, Sheraton
I-496 & Creyts Road,
Lansing, Michigan



Dr. Vernon Hoffner, CTO
EntreSoft Resources, Inc

Dr. Vernon Hoffner will be the keynote speaker at the DAMA Michigan Program to be held on January 8, 2002.

For the past 15 years Vernon Hoffner has been involved in Senior Level IS projects, providing both management and technical leadership. He has 35 years of experience in the Information Systems and Management Science fields, including positions of programmer, systems analyst, and manager. Currently, Chief Technology Officer at EntreSoft Resources, Inc., he is responsible for business development in the areas of IS management and decision support. His areas of strength and interest include: Management Problem Solving, Strategic and Tactical IS Planning, Project Management, and Information Sys-

Data Warehousing 101

Data warehousing has been around for over a decade. Therefore, when you read the articles in professional magazines and journals, most authors assume that you know and understand the terminology and acronyms. This can be a problem when you are now trying to understand this area of the technology. We will review the technology and terminology (buzzwords) and provides an integrated foundation for understanding the benefits and potential hazards for data warehousing. We will begin with some definitions and their implications for understanding data warehousing. Next, is a discussion of how data warehousing fits within the IT infrastructure, its benefits and application.

Larry English Presents “Elevating the Role of Information Resource Management for Business Effectiveness” February 12, 2002 at Novi Doubletree

Larry English, President of Information Impact International, Inc. will present his program on *Elevating the Role of Information Resource Management for Business Effectiveness* to the member-

ship of the DAMA Michigan Chapter of DAMA International at the Novi Doubletree Hotel from 11:30 to 4:30 on February 12, 2002. Reservations for the event can be made by contacting Michelle Rymer at (248) 265-9355 or

completing the registration form available on the Chapter web site at dama-michigan.org
An abstract of Mr. English’s presentation follows on page 3.

(Continued on page 3)

Point of View

Point of View

By W. Thomas Hamlin

Comments: thomashamlin@hotmail.com

Dear Reader,

In the last issue, I offered a suggested topic for this issue, entitled The Next Y2K Problem. O.K., I was intentionally overdramatic. The topic I am writing about today does not offer all the drama that Y2K did, but nonetheless it provides room for improvement in our data design efforts.

I am suggesting that with a little more effort and analysis of common data values, we can greatly expand the downstream functionality of our

we can greatly expand the downstream functionality of our systems and add value at the same time without adding appreciable cost to any project

systems and add value at the same time without adding appreciable cost to any project. We need to use some forward thinking about our data designs and keep in mind other potential uses of our data.

To the point, major challenges for IT are customer-oriented (CRM, B2B, B2C) systems that can provide a platform for in depth customer analysis by the firm. For example, if you have or

are designing a system that captures customer information such as name, address, phone number and social secu-

urity number you can be sure an analyst will want to 'slice and dice' the information by those values. For example, 'give me all customers by street name with this ZIP code'. Don't try to understand why anyone would want this information, they just do! What's more, we should be able to provide it.

Review the following common examples:

1. 1234 E. Main Street, S.W.
2. 48910-1234
3. J. Albert Smith
4. (515) 888-6194

Internationally Recognized Speaker Larry P. English Book Signing February 12

Larry P. English, president and principal of INFORMATION IMPACT International, Inc., is an internationally recognized speaker, teacher, consultant, and author in information quality improvement. He has provided consulting and education in more than 25 countries on five continents. He was featured as one of the "21 Voices for the 21st Century" in the January, 2000 issue of *Quality Progress*. DAMA awarded him the 1998 "Individual Achievement Award" for his contributions to the field of information resource management. He has organized and chaired 8 Information and Data Quality Conferences in the US and Europe since 1997.

Mr. English's methodology for information quality improvement—Total Quality data Management (TQdM®)—has been implemented in several organizations worldwide. He writes the "Plain English on Data Quality" column in the *DM Review*. Mr. English's widely acclaimed book *Improving Data Warehouse and Business Information Quality*, has been translated into Japanese.

2002 Membership Drive

By Carl Guse
Vice President - Membership

Membership registration for 2002 is getting underway and we are soliciting the help of our current membership to identify new prospects to join our DAMA-Michigan chapter.

Currently we have 133 members carrying over from 2001 to 2002. Those of you who signed up in 2001 took advantage of our 2 year for 1 year special so your registration dues are already covered for 2002.

At a recent Board Meeting we established a goal of growing our membership by 10% per year. However, looking at our current list of members and member companies, I know we have only touched the tip of the iceberg relative to the number of potential members we could have here in Michigan. My hope is that we can grow our membership by 20% or better for 2002.

So, what we are asking our current membership to do is to look within and outside of your current organization to see what other Data Management staff you interact with and send us one name and contact information

for that person so that we can invite them to join our Chapter. Better yet, would be if you invited them directly; but I will gladly do this for you if you can get me the contact information I need.

Corporate Sponsors that joined in 2001 and carry over for 2002 are:

General Motors
EDS
Universal Solutions, Inc.
Innovative IT Solutions

Corporate Membership companies that also carry over through 2002 are:

AAA Auto Club of Michigan
Allmerica Financial -
Citizen's Insurance
Burnett Direct Inc.
DOW Chemical
Mount Clemens General Hospital
Proforma Corporation
Blue Cross Blue Shield of Michigan
Kmart
State of Michigan -
Dept. of Management & Budget
State of Michigan -
Dept. of Transportation
State of Michigan -
Office of Retirement Services

(Continued on page 5)

The following is an abstract of the program to be presented on February 12, 2002 by Larry English. There will be a discussion panel and book signing following the presentation.

The organization that is not managing its information cannot manage its business. Without managed, quality information, the enterprise cannot "know" what it needs to know to understand its customers and customer needs, manage operations, analyze its performance and make the strategic decisions for the future of the enterprise. This is even more crucial for service sector organizations, such as banks, insurance and government organizations whose products are, in fact, information.

Mr. English describes how you transform and elevate your data administration or data resource management to a value-adding function embraced as a critical business competency. The principles used to manage other business resources, such as human and financial resources, apply to managing information and knowledge as strategic resources. Implementation of these principles is required to transform the enterprise from an Industrial-Age to a competitive Information-Age organization. This presentation discusses how the organization can harness the power of today's information technology to exploit its information resources for competitive advantage and business effectiveness.

- Why traditional approaches to data administration have failed to create positive impact and acceptance in the enterprise
- Why the "systems approach" of application development has failed, and how we must replace it
- Trends shaping the economy, business, and society; and its impact on information resource management
- The Information Age as a paradigm
- From data administration to information stewardship
- Information Resource Management in the e-business world: the virtual enterprise
- The secrets to gaining and sustaining management "commitment"

Larry P. English Educator and Author to Address DAMA Michigan

Mr. English is an internationally recognized speaker, educator, author and consultant in information and knowl-



edge management and information quality improvement. He also provides consulting and education in information stewardship, strategic information visioning, information technology evaluation, information resource management and data administration, data modeling and facilitation, and value-centric application development methods. Mr. English has developed the Total Quality data Management (TQdM[®]) methodology applying Kaizen[®] quality principles to information quality management. He chairs Infor-

mation Quality Conferences around the world.

Prior to founding INFORMATION IMPACT INTERNATIONAL, Inc. (www.infoimpact.com), Brentwood, TN, over twelve years ago, Mr. English was Vice President of an international IRM consulting firm. Before that, he was manager of systems development and then for information management with a large publishing firm. Before positions as Senior Instructor for a computer manufacturer and Information Systems Training Coordinator for a major insurance firm, Mr. English began his career with Sears, Roebuck, and Co., as a programmer and systems analyst.

He was featured as one of the "21 Voices for the 21st Century" in the January, 2000 issue of *Quality Progress*. DAMA awarded him the 1998 "Individual Achievement Award" for his contributions to the field of information resource management. Mr. English has served as an Adjunct Associate Professor in computer science. Active in several professional organizations, he has been an officer of the

Nashville DPMA Chapter and is a co-founder of the Nashville DAMA Chapter. He is a member and a strategic business partner of the American Society for Quality (ASQ). Mr. English has been an active member of various ANSI (American National Standards Institute) standards committees, and he is editorial advisor for *DM Review*.

A magna cum laude graduate of Hardin-Simmons University, Mr. English holds a Masters Degree from the Southern Baptist Theological Seminary where he was a Luther Rice Scholar and a Garrett Fellow. He is listed in Outstanding Young Men in America and Who's Who Worldwide. He has provided consulting and educational services in more than 25 countries on five continents to such organizations as Aera Energy, Air Canada, American Express, Belgacom, Boeing, British Telecom, Capital Bank, Coca-Cola Foods, Dow Chemical, Eastman Kodak, Eli Lilly, the FDIC, Hewlett-Packard, The Hartford, IBM, L. L. Bean, NTT DATA, Optical Fibres, Sprint, Telenor, UNUM Life Insurance

(Continued on page 7)

Information Management as a Business Function

by Nancy K. Dwyer Many Data Management organizations have emphasized the importance of managing data as a corporate resource, but how many companies have really implemented a business function to do so? If managing data and information is truly as important as managing human, financial, and material resources, why is Information Management (IM) often not identified as a legitimate business function, with its own processes, data and information systems to support it?

Many organizations primarily produce INFORMATION of some kind as their “product.” Even in manufacturing companies, many of the employees produce information which allows their factories to build and ship their products. Much of their cost of doing business is the costs associated with these “information workers”. Doesn’t it make sense to manage and optimize the producing of this information?

I believe that IM is a legitimate business function concerned with identifying, documenting, automating and managing these processes and data in order to support company operations, and implementing them in an appropriate information systems infrastructure.

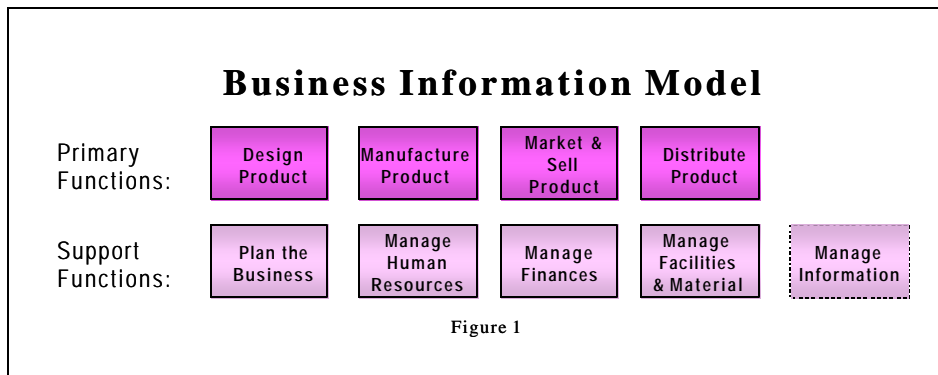
Many companies use a Business Information Model (BIM) to define the major business functions of the company. As in the example in Figure 1, a BIM is often divided into primary functions: the functions needed to develop and deliver the products and/or services of the company; and the support functions: functions which the company needs to perform in order to support the execution of the primary functions.

Data Modeling Notation Extended

By Jerry E. Stembridge
VP Online Communications

Since the introduction and standardization of the UML (Unified Modeling Language) by the OMG (Object Management Group) some data modeling products have been extended to offer data modeling using the UML notation as opposed to the traditional notations like IDEF1X and Information Engineering (IE). The table on page 6 provides a brief description of the data modeling concepts used in IDEF1x and describes how these same concepts can be modeled using the UML Class Diagram. An expanded version of this comparison can be found on our WEB site at www.dama-michigan.org.

(Continued on Page 6)



The BIM serves as a framework to document their business processes. Associated with these business processes is the definition of the data, or metadata, that the company needs to execute the business processes, and the data that is produced by these processes to be used by downstream processes. These processes need to be documented in process models, and the data used by each process standardized, documented in data models, and associated to the processes that use it. Automatable processes should be implemented in accordance with the process models in common application systems, used by every division/department that needs to perform a given business process. The data to support these common applications should be implemented in accordance with the data models in sharable data bases, accessible to all applications which need to use the data. The common applications and sharable data bases should be implemented using an appropriate technology infrastructure, based on an enterprise-level technology architecture designed to meet the company’s needs.

The IM business function needs to manage the content of each box in Figure 2, and manage the associations between them indicated by the arrows. The logical components, including the processes, data and technology architecture, should be documented in the form of models and stored in an appropriate repository. These model

(Continued on page 7)

What are 8 good reasons to join DAMA Michigan?

1. Professional Programs
2. Professional Networking
3. Education Programs
4. National and International Speakers
5. Resource Library
6. Professional Collaboration
7. Compare Notes with other DM Professionals
8. Enjoy a Pleasant Meal with Colleagues

Join Today!
Contact Carl Guse
at 248-945-5527 ext 527
or
Go directly to our website
www.dama-michigan.org

2002 Membership Drive

(Continued from page 2)

With exciting programs developing for 2002 (i.e., Larry English in February and Claudia Imhoff in May) it is important that new members send in their Membership Application and payments early in order to receive the Membership pricing for these events. Members signing up after February can pay a prorated amount instead of the full amount of annual dues.

Membership levels and pricing for 2002 are as follow:

1. Corporate Sponsor

Dues: \$500

Benefits:

- No limit on number of people who can attend regular DAMA meetings
- No limit on number of people who can attend special DAMA-Michigan events at member rate
- Designated members receive DAMA-Michigan newsletter
- Identification as a DAMA-Michigan Sponsor on newslet-

ters and web site

- Opportunity of hosting regular DAMA speaker meetings at company's headquarters
- One Ballot for official Chapter voting processes

2. Corporate Member

Dues:

Level 1: 1-5 people: \$150

Level 2: 6-10 people: \$250

Level 3: 11+ people \$350

Benefits:

- Designated members can attend regular meetings
- Designated members can attend special DAMA-Michigan events at member rate
- Designated members receive DAMA-Michigan newsletter
- Identification as a DAMA-Michigan Corporate Member on newsletters and web site
- One Ballot for official Chapter voting processes

3. Individual Member

Dues: \$45

Benefits:

- Member can attend regular meetings
- Member can attend special DAMA-Michigan events at member rate
- Member receives DAMA-Michigan newsletter
- One Ballot for official Chapter voting processes

4. Student Member

Dues: \$15

Benefits:

- Member can attend regular meetings
- Member can attend special DAMA Michigan events at member rate
- Member receives DAMA-Michigan newsletter
- Has no Chapter voting rights

DAMA Chapters around the globe have to rely on the dues of it's membership as it's financial lifeline. The larger our paid membership base is, the more exciting events we can plan. Feedback

(Continued on page 7)

Point of View

(Continued from page 2)

5. 388-94-7891

Simplistic textbook examples would encourage the data analyst to reach the following conclusions, 1 is an address, 2 is a ZIP code, 3 is a name, 4 is a phone number and 5 is a Social Security Number (SSN). None of the above examples are atomic or scalar in nature. That is, each contains multiple data values. This simply is a violation of sound relational database design.

Worst yet the DBA implements this logical model as a physical one. This is a 'D minus' solution. It does provide some functionality but it creates many more problems that it solves for down stream analytics. String manipulation functions will parse out the street name etc, but it all takes lots of CPU time,

sophisticated coding and it is not at all necessary.

Below I am offering the following alternatives for your consideration. Column names are only suggestions, but it makes sense to me now and hopefully to anyone else accessing the data in the future. Does anyone really know here the data dictionary is?

My suggested solutions,

- 1) Address (at least 4 different data elements), i.e., Street Number, Street Directional, Street Name and City Quadrant. Now searching for 'Main Street' is very easy. I am not convinced that potentially separating StreetName from StreetType (Ave, Blvd, St, etc) is worth the effort or adds

value. Yes, I know I already violated the atomic data value rule.

- 2) ZIP Code (2 different data elements), i.e., ZIP Code and ZIP Plus Four. Again, searching is much easier.
- 3) Name (at least 3 different data elements and optionally two more), i.e., First Name, Middle Name, and Last Name. We hit the jackpot here. By adding two additional data elements, Name Prefix, we can help identify gender and with Name Suffix, we can help separate cohort/generations. I would create an application design using dropdown lists with predetermined

(Continued on page 7)

Comparison of OO's UML vs ER's IDEF1x. Modeling Languages

By Jerry E. Stembridge VP Online Communications

Unified Modeling Language (UML)	Integrated Computer Aided Manufacturing (ICAM) Definition Language (IDEF1x)
<p>Class - In the UML the Class Diagram is used to model data requirements. A UML class is modeled as a box.</p> <p>The major difference between a UML class and an IDEF1x entity is the presence of methods in a class. A method represents a service that can be requested of a class to affect its behavior (i.e., perform operations on its attributes). An IDEF1x entity can be thought of as a UML class without operations.</p>	<p>Entity - An entity represents a set of "things" such as persons, places, abstractions, or things that are relevant to a business and about which data is collected and maintained. These things share a set of characteristics by which they can be uniquely described. An individual member of the set is referred to as an entity instance. An IDEF1x entity is modeled as a box.</p>
<p>Association - In the UML a relationship is called an association. An association is modeled using a solid line that connects two classes.</p>	<p>Relationship - A relationship is an association between two or more entities that reflects a relevant business rule. In IDEF1x a relationship is modeled as a line drawn between entity boxes, with a dark circle on one end of the line. The entity touched by the dark circle is called the child entity and the entity at the other end of the line is called the parent entity. The relationship lines differ in appearance based on the idea of primary key inheritance.</p>
<p>Multiplicity - In the UML cardinality is called multiplicity. Multiplicity refers to the number of objects in one class that can be related to one object in the related class.</p>	<p>Cardinality - In an IDEF1x data model relationships between entities are constrained by adornments at each end of the relationship line. These constraints are called cardinality. Cardinality represents the maximum number of times an entity instance in one entity can be related to the set of entity instances in another entity.</p>
<p>Associative Class - In the UML an IDEF1x dependent entity is modeled with an association class. The associative class is modeled as a class (box) connected to an association line (that is a solid line), with a dashed line.</p>	<p>Dependent Entity - In IDEF1x the child entity in a relationship is said to be existence dependent upon the parent entity in that an instance of the child entity can only exist if the associated instance of the parent entity exists. For example, in a relationship between the entities <i>customer</i> and <i>policy</i>, a customer is the insured-party for zero, one, or more policies and each policy must insure exactly one person.</p>
<p>N-ary Association - An N-ary association is an association between three or more classes. In the UML an N-ary association is modeled by connecting each class that participates in the N-ary association to a large diamond using an association line (that is a solid line).</p>	<p>Dependent Entity - In IDEF1x an N-ary relationship is modeled using a dependent entity where the dependent entity is dependent upon all of the entities that participate in the relationship.</p>
<p>Subtype - In the UML an IDEF1x category is modeled as a subtype class. In both the UML and IDEF1x the supertype of a set of subtypes is referred to as a generalization and the constructs used to represent a subtype / supertype relationship is referred to as a generalization hierarchy. The child or subtype in the hierarchy inherits all of the characteristics of the parent or supertype. In the UML a subtype / supertype association is modeled by connecting each subtype class to the supertype class with a block arrowhead</p>	<p>Category - Entities represent things about which we need to collect and keep information. Some entities are categories of other entities. For example a bank collects and keeps information about customer accounts. Some accounts are checking and some are savings. Some of the information collected about accounts is applicable to both checking and saving accounts, e.g., owners name and address. However, some information is unique to either checking or savings accounts. Therefore, the entities <i>saving-account</i>, and <i>checking-account</i> are categories of the entity <i>account</i>. In IDEF1x these entities are modeled using a categorization relationship, which is also called a subtype, or an "is-a" relationship.</p>

INFORMATION RESOURCE MANAGEMENT

(Continued from page 4)

should be used as a Master Plan; any system development project must implement a piece of the Master Plan in order to be justified. There should be standards for the technology infrastructure products that will be used to implement the technology architecture components, and these standard products should play well together.

Defining IM as a legitimate business function and putting in place the processes, data, infrastructure and funding to support it, is a major step controlling a company's information systems environment. Figure 3 (shown on page 8) contains an IM Business Model that can be used as a starting point for identifying IM business processes.

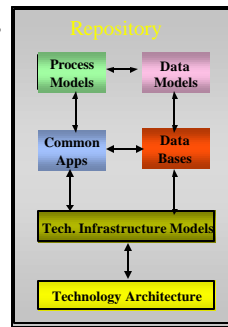


Figure 2

(Continued on Page 8)

2002 Membership

(Continued from page 5)

we have gotten from the programs in 2001 indicate that our meetings are providing our Michigan members opportunities to network with their peers from other companies and this is one of the important aspects of being a DAMA-Michigan member.

As a current officer of the DAMA-Michigan Board, I still feel that the

Michigan Chapter can develop into one of the larger chapters within this region and would love to see our membership grow to being one of the largest in the US.

I am available by phone at (248)945-5527 or email at carl@usi-online.com for your comments or questions regarding membership issues.

Larry English, Author and Educator

(Continued from page 3)

Co., the U.S. Navy, Western Health Alliance and Weyerhaeuser.

A frequent keynote speaker, Mr. English writes the monthly "Plain English on Data Quality" column for *DM Re-*

view, and is the author of the highly acclaimed *Improving Data Warehouse and Business Information Quality*, now available in Japanese, and numerous articles for publications in the US and Europe.

Point of View

(Continued from page 5)

values to ensure consistency. In addition, this will certainly help with names of persons with different ethnic backgrounds.

4) Phone Number (3 data elements), i.e., Phone Area Code, Phone Exchange, and finally Phone Number. Now if you need to change the PhoneAreaCode value, while still not simple, it is much easier.

5) Social Security Number (only 1 legitimate value, if that). Yes, each set of numbers does have a specific purpose/meaning, but unless you are the Social Security Commission, I would suggest it is none of your organization's business. This is potentially a very explosive/emotional/security/privacy laden issue. Further, unless you have a legitimate reason for capturing this information, I would stay as far away from SSN as I could.

Fess up, how many of your systems actually use SSNs as primary keys? How difficult will it be to modify your systems, if required? Sound like Y2K size problems to me! I would get started now on a solution.

Next Issue: Data Warehouse or Junk Yard. Good design is a choice.



Dr. Vernon Hoffner
Data Warehousing 101
Reserve Now!
Limited Seating!
Lansing, Sheraton
January 8, 2002
5:30 p.m.-8:30 p.m.
Dinner and Program
www.dama-michigan.org

Coming Events

Meeting Date	Meeting Time	Guest Speaker	Topic	Location
January 8, 2002 Lansing, Sheraton	5:30—8:30	Dr. Vernon Hoffner, CCP EntreSoft Resources, Inc.	Data Mining	TBD
February 12, 2002 Novi, Doubletree	11:30 —4:30	Larry English Information Impact Int., Inc.	Information Resource Management for Busi- ness Effectiveness	TBD
March 12, 2002	5:30 - 8:30 .	TBD	TBD	TBD
April 9, 2002	5:30 - 8:30	Joe Oates Sybase, Inc.	TBD	TBD
May 17, 2002	5:30 - 8:30	Claudia Imhoff Intelligent Solutions, Inc	TBD	TBD
June 12, 2002	5:30—8:30	TBD	TBD	TBD

Information Management

(Continued from page 7)

Figure 3

Technical

- Maintain information about the Technology Infrastructure used by current common Applications and shared Data Bases
 - Collaborate with Process Management in maintaining mappings between information about common Applications and Technology Infrastructure
 - Collaborate with Data Management in maintaining mappings between information about Data Bases and Technology Infrastructure
 - Define standard Technology Architecture (s)
 - Maintain mappings of Technology Infrastructure to Technology Architecture(s) they implement
 - Maintain standards for Products to implement standard Technology Architectures
- Manage compliance to technology standards and direction

Business

- Maintain the Business Model
- Maintain the mapping between the Business Model and the Process Models
- Maintain the Process Models
- Collaborate with Data Management in maintaining mappings between Process Models and Data Models
- Manage common Application development and maintenance
- Manage compliance to System Development Processes

Technical

- Maintain information about common Applications
 - Maintain mapping between Process Models and common Applications
 - Collaborate with Data Management in maintaining mappings between information about common Applications and sharable Data Bases
 - Collaborate with Technology Management in maintaining mappings between information about common Applications and Technology Infrastructure elements
 - Develop, operate and maintain common Applications
- Develop and maintain sharable Basic Processes for use in Applications

Business

- Develop and maintain Data Models to document Corporate Data Standards
- Collaborate with Process Management in maintaining mapping between Process Models and Data Models
- Manage shared Data Base development and maintenance
- Manage compliance to Data Resource Management processes

Technical

- Maintain information about shared Data Bases
 - Maintain mapping between Corporate Data Standards Data Models and Data Bases
 - Collaborate with Process Management to maintain mapping between Data Bases and common Applications
 - Collaborate with Technology Management to maintain mappings between information about shared Data Bases and Technology Infrastructure elements
 - Develop, operate and maintain shared Data Bases based on Data Models
 - Provide common Applications access to shared data
- Coordinate maintenance information in an Information Management Repository