Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be “used for any purpose other than private study, scholarship, or research.” If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of “fair use” that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select “Pages from: first page # to: last page #” on the print dialog screen
The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.
ABSTRACT

TOWARDS HYPERMEDIA SUPPORT IN DATABASE SYSTEMS

by

Anirban Bhaumik

The general goal of our research is to automatically generate links and other hypermedia related services to analytical applications. Using a dynamic hypermedia engine (DHE), the following features have been automated for database systems. Based on the database's relational (physical) schema and its original (non-normalized) entity-relationship specification links are generated, database application developers may also specify the relationship between different classes of database elements. These elements can be controlled by the same or different database application, or even by another software system. A DHE prototype has been developed and illustrates the above for a relational database management system.

The DHE is the only approach to automated linking that specializes in adding a hyperlinks automatically to analytical applications that generate their displays dynamically (e.g., as the result of a user query). The DHE's linking is based on the structure of the application, not keyword search or lexical analysis based on the display values within its screens and documents. The DHE aims to provide hypermedia functionality without altering applications by building "application wrappers" as an intermediary between the applications and the engine.
TOWARDS HYPERMEDIA SUPPORT IN DATABASE SYSTEMS

by
Anirban Bhaumik

A Thesis
Submitted to the Faculty of New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science

Department of Computer and Information Science

January 2001
TOWARDS HYPERMEDIA SUPPORT IN DATABASE SYSTEMS

Anirban Bhaumik

Dr. Michael Bieber
Associate Professor of Computer and Information Science, NJIT

Dr. Vincent Oria
Assistant Professor of Computer and Information Science, NJIT

Dr. Byoung-Kee Yi
Assistant Professor of Computer and Information Science, NJIT
BIOGRAPHICAL SKETCH

Author: Anirban Bhaumik
Degree: Master of Science in Computer Science
Date: January 2001

Undergraduate and Graduate Education:

• Master of Science in Computer Science,
  New Jersey Institute of Technology, Newark, NJ 2001

• Master of Science in Chemical Engineering
  New Jersey Institute of Technology, Newark, NJ, 1998

• Bachelor of Engineering in Chemical Engineering
  Regional Engineering College, Durgapur, India, 1996

Major: Computer Science

Presentations and Publications:

“Computer Aided Cognitive Tools for Teaching and Implementing Clean Manufacturing” at
the National Science Foundation, Technology Reinvestment Project, Engineering Education
Innovators’ Conference, April 8, 1997, Washington DC.
To my Parents, and Suchi, thanks for being there.
ACKNOWLEDGMENT

The author would like to thank his advisor Dr. Michael Bieber for his guidance. Special thanks are also due to my colleagues in the Collaborative Hypermedia Laboratory at the New Jersey Institute of Technology: Mr. Roberto Galnares, Mr. Firas Aljallad, Ms. Deepti Dixit and Ms. Aparna Reddy.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Motivation</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Literature Review</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1 Hypermedia Engines</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2 Research in Hypermedia and Databases</td>
<td>3</td>
</tr>
<tr>
<td>2 The Dynamic Hypermedia Engine</td>
<td>7</td>
</tr>
<tr>
<td>3 Hypermedia support for relational database management systems</td>
<td>12</td>
</tr>
<tr>
<td>3.1 Introduction to the Relational Database Wrapper Module</td>
<td>12</td>
</tr>
<tr>
<td>3.2 Elements of Interest in Database Systems</td>
<td>14</td>
</tr>
<tr>
<td>3.2.1 Element Types</td>
<td>14</td>
</tr>
<tr>
<td>3.2.2 Marking an Element of Interest</td>
<td>15</td>
</tr>
<tr>
<td>3.2.3 Element Identifiers</td>
<td>16</td>
</tr>
<tr>
<td>3.2.4 Marked up Message</td>
<td>16</td>
</tr>
<tr>
<td>3.3 Mapping Rules and Metadata for Relational Databases</td>
<td>18</td>
</tr>
<tr>
<td>3.3.1 Mapping Rules for Relational Database Management Systems</td>
<td>19</td>
</tr>
<tr>
<td>3.3.2 Metadata for Relational Database Management Systems</td>
<td>20</td>
</tr>
<tr>
<td>3.4 The Relational Database Wrapper Module</td>
<td>28</td>
</tr>
<tr>
<td>3.4.1 Features</td>
<td>28</td>
</tr>
<tr>
<td>3.4.2 Object Oriented Analysis</td>
<td>41</td>
</tr>
<tr>
<td>3.4.3 Functionality</td>
<td>42</td>
</tr>
<tr>
<td>3.5 Enhanced Links through a Database Schema Wrapper</td>
<td>50</td>
</tr>
<tr>
<td>4 Providing Support to Database Applications</td>
<td>55</td>
</tr>
</tbody>
</table>
## TABLE OF CONTENTS
(Continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Support to existing Database Applications.</td>
<td>55</td>
</tr>
<tr>
<td>4.1.1 Motivation</td>
<td>55</td>
</tr>
<tr>
<td>4.1.2 Architecture</td>
<td>55</td>
</tr>
<tr>
<td>4.1.3 Case Study</td>
<td>57</td>
</tr>
<tr>
<td>4.2 Support to Applications being developed</td>
<td>58</td>
</tr>
<tr>
<td>4.2.1 Motivation</td>
<td>58</td>
</tr>
<tr>
<td>4.2.2 Architecture</td>
<td>59</td>
</tr>
<tr>
<td>4.2.3 Case Study</td>
<td>61</td>
</tr>
<tr>
<td>4.2.4 Advantages of using the DHE’s infrastructure</td>
<td>62</td>
</tr>
<tr>
<td>5 Using the DHE as a Data Warehousing Solution.</td>
<td>64</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>64</td>
</tr>
<tr>
<td>5.1.1 Data Warehouse Metadata</td>
<td>64</td>
</tr>
<tr>
<td>5.1.2 Motivation</td>
<td>65</td>
</tr>
<tr>
<td>5.2 Functionality</td>
<td>65</td>
</tr>
<tr>
<td>5.2.1 The Loader Module</td>
<td>65</td>
</tr>
<tr>
<td>6 Future Research Plans – integrating data mining</td>
<td>68</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>68</td>
</tr>
<tr>
<td>6.2 Mapping Rules for Data Mining</td>
<td>68</td>
</tr>
<tr>
<td>6.2.1 Two Step Mapping Rules</td>
<td>69</td>
</tr>
<tr>
<td>APPENDIX A RDF Resource Properties used by the RDWM</td>
<td>70</td>
</tr>
<tr>
<td>APPENDIX B Database Connection Pool Performance Metrics</td>
<td>78</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>80</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Screenshot of the DHE</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>DHE Architecture</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Relational Database Wrapper Module Message Flow</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>Marked up Message</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>Metadata of a Column</td>
<td>22</td>
</tr>
<tr>
<td>7</td>
<td>RDF/XML Syntax</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>Class diagram of the “dhyme.metadata.rdf” package</td>
<td>24</td>
</tr>
<tr>
<td>9</td>
<td>Class diagram of the “dhyme.metadata.database” package</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>RDWM State Diagram</td>
<td>29</td>
</tr>
<tr>
<td>11</td>
<td>Connection Pool Architecture</td>
<td>35</td>
</tr>
<tr>
<td>12</td>
<td>State Diagram for Requesting and Returning Connections</td>
<td>37</td>
</tr>
<tr>
<td>13</td>
<td>Broker Housekeeping State Diagram</td>
<td>38</td>
</tr>
<tr>
<td>14</td>
<td>Class Diagram of the dhyme.utils.dbBroker.DBBroker Class</td>
<td>39</td>
</tr>
<tr>
<td>15</td>
<td>RDWM Class Diagram</td>
<td>41</td>
</tr>
<tr>
<td>16</td>
<td>Displaying the UI - Collaboration Diagram</td>
<td>43</td>
</tr>
<tr>
<td>17</td>
<td>Executing a select statement - Collaboration Diagram</td>
<td>44</td>
</tr>
<tr>
<td>18</td>
<td>Executing a non-select statement - Collaboration Diagram</td>
<td>45</td>
</tr>
<tr>
<td>19</td>
<td>Retrieving Primary Keys</td>
<td>46</td>
</tr>
<tr>
<td>20</td>
<td>Retrieving Foreign Keys</td>
<td>47</td>
</tr>
<tr>
<td>21</td>
<td>Retrieving a Table</td>
<td>48</td>
</tr>
<tr>
<td>22</td>
<td>Retrieving a Tuple</td>
<td>49</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>23</td>
<td>Retrieving Metadata</td>
<td>50</td>
</tr>
<tr>
<td>24</td>
<td>Screenshot of the DSMM</td>
<td>51</td>
</tr>
<tr>
<td>25</td>
<td>Providing Support to New Database Applications – Architecture</td>
<td>59</td>
</tr>
<tr>
<td>26</td>
<td>State Diagram</td>
<td>60</td>
</tr>
<tr>
<td>27</td>
<td>DHE Data Warehouse Loader Module</td>
<td>67</td>
</tr>
<tr>
<td>28</td>
<td>Database Connection Pool Performance at a Glance</td>
<td>78</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Motivation

Database queries typically return results in a plain text format. Some applications on the World Wide Web generate link anchors for database elements, but these anchors normally hold a single link to the most obvious destination for the dominant type of user.

An element within a database application maybe considered a potential starting point for information exploration. Each element may have multiple links, each representing a different relationship (schema-based or otherwise). The ability to explore a piece of information in more detail would allow users to better understand that item, as well as analyze and view the various relationships that define these elements. Users may wish to explore around data values and symbols they see, labels on graphs or user input forms, options in pop-up lists, or even on the menu commands they can invoke.

To complicate the developer's job, users often have different mental models of an application and its underlying domain than the developer. Even when developers work closely with users, the end result might not be intuitive for all users or serve each user's individual tasks equally well. Many people visit a given application's screen aside from the most dominant type of user(s) for which it was developed. These include other users of the application, customer service representatives, company analysts, managers, trainees, people inside the company designing new databases or applications based on the current one, external analysts, and stockholders, among others. Each may be interested in different aspects of application elements, according to their current task-at-hand. Customization is one solution, but even so users often might wish to explore several different relationships from a given anchor, and therefore should have several links available.

The purpose of this research is to explore all aspects of hypermedia support for database applications, and is based on the experience of designing and developing the prototype Dynamic Hypermedia Engine (DHE). The DHE automatically generates anchors, sets of links and metadata within database applications, as well as supporting users with other types of
hypermedia structuring, navigation and annotation functionality, including guided tours and annotation.

This work makes many contributions to both the database and hypermedia fields. Many database applications do not take as much advantage of hypermedia as they could. This chapter puts forth a series of opportunities for integrating hypermedia and database systems. As we shall describe in the next chapter, the DHE is the only tool that provides automated linking and hypermedia services based on the application structure (as opposed to search or lexical analysis), without altering applications. Thus it is uniquely suited to support databases and other analytical applications on the Web that generate the contents of their displays dynamically in response to user queries.

This work proceeds as follows. A Literature Review of other Hypermedia Engines and hypermedia support in databases is presented followed by an introduction the Dynamic Hypermedia Engine (DHE). The following chapters show how the DHE provides support to relational DBMS, database applications and enterprise-wide Data Warehouses. This research concludes by providing a direction for future Research in this field.

1.2 Literature Review

1.2.1 Hypermedia Engines
Several approaches exist for integrating hypermedia functionality into primarily non-hypermedia information systems. These include employing hypermedia data models (Campbell and Goodman 1988; Halasz and Schwartz 1994), hypermedia toolkits (Anderson 1996), link services (Pearl 1989; Davis et al. 1992; Anderson 1997), hyperbases (Leggett and Schnase 1994), hypermedia development environments (Nanard and Nanard 1995a; Marshall and Shipman 1995; Akscyn et al. 1988), open hypermedia systems (Whitehead 1997; Wiil 1997; Grønbæk and Trigg 1999), and independently executing hypermedia engines, such as the DHE.

Hypermedia engines execute independently of an application with minimal modifications to it, and provide the application's users with hypermedia support. Few approaches provide transparent hypermedia integration as our engine does. Notable projects include Microcosm's Universal Viewer, Freckles and the OO-Navigator and SFX.
Microcosm's Universal Viewer (Davis et al. 1994) and Freckles (Kacmar 1993, 1995) seamlessly supports an application's other functionality but provides only manual linking. OO-Navigator comes the closest to our approach, providing a seamless hypermedia support for computational systems that execute within a single Smalltalk environment (Garrido and Rossi 1996; Rossi et al. 1996). This approach meets our goal of supplementing Smalltalk applications with hypermedia support without altering them. Our approach applies to both object-oriented and non object-oriented applications.

SFX's engine is very similar to DHE, but it only serves one specific environment. SFX dynamically generates anchors within the reference section of academic papers being displayed on the Web. Selecting these will lead to the original work within bibliographic databases (Van der Stemple, 1999a,b,c). DHE, in contrast, provides a generalized approach for linking and additional hypermedia functionality for most analytical applications.

1.2.2 Research in Hypermedia and Databases.
Several techniques have been proposed recently for the integration of hypertext and databases. Some of them address the issue of building hypertext structures over existing databases to provide more direct navigation through hyperlinks.

Hara and Botafogo (Hara et al 1994) use an SQL-like data definition language to map single relations or relational views to node types. A node type is similar in nature to an entity type, i.e., it models a real world object or concept in the hypertext conceptual schema defined over the database contents. Its specification includes the correspondences between relation attributes and node fields, as well as presentation information. At run time, a node type produces two kinds of nodes: a composite one for the whole relation, and a number of nodes corresponding to the tuples of the relation. The same language is used to define link types among node types. A WHERE-clause is used to constraint the creation of links during navigation.

In a similar approach, Falquet et al. (Falquet et al 1995, Falquet et al 1998) offer a declarative language to produce databases views composed of node and link schemas, accessed through the WWW. Each node schema is based on one object class or a set of inter-related object classes. The content of the node is composed of a subset of the attributes of the class(es).
Foreign keys to other classes constitute link types to the corresponding nodes. Two kinds of links are supported: *Reference* links are indented to offer navigation structure within the nodes, while *inclusion* links are indented to create nested structures (part-of relationships). In addition, the specification of the relational view includes presentation information. The above definitions form the input to a cgi-script that produces HTML pages for the end-user. The DHE would enhance the existing views specified through the database application.

The above approaches leave the original client application intact, introducing a new interface that provides hypermedia-based interaction with the database. On the other hand the DHE overlays linking facilities within the original user interface application by means of user interface wrappers.

*Domenicus* (Constantopoulos et al. 1996) is a hypermedia engine developed over a repository management system, called Semantic Index System (SIS). Domenicus offers hypermedia functionality (such as alphabetic lists, subject catalogs, guided tours, query cards, hyperlinks, image annotations, bookmarks and history), based on predefined queries over the information objects and their structure, managed by SIS. *Presentation Card Specifications* are executed at runtime to present objects or classes of objects stored in SIS, while *hyperlink classes* dynamically produce links during navigation. Different presentation models can coexist for the same repository instance, to fulfil the searching, browsing and updating requirements of different user groups. The DHE provides many of these features in a generic way over any application, allowing tours and indexes to contain elements from several systems. Also, in the DHE queries are only predefined to the extent that mapping rules hold skeleton queries for particular classes of database elements.

Other approaches suggest embedding database queries into HTML pages. For example, a mechanism offering cross-language variable substitution between HTML and SQL is the core of the *DB2 WWW Connection* system (Nguyen et al. 1996), which enables quick and easy construction of applications accessing relational databases from the Web. The developer creates *macros* that consist of HTML and SQL commands, written in distinct sections. The sections are tied together via variable substitution. Macros are stored at the Web server and are processed by cgi-scripts in order to get user input or produce output reports. The DHE
does not store single database queries in the pages displayed on the Web. Instead we generate a list of several possible links for any element from specifications in the mapping rules.

Instead of providing hypertext functionality for a specific database, Geldof (Geldof 1996) uses an abstract page definition language to construct templates embodying presentation guidelines for terms of an ontology; a conceptualisation containing objects, concepts and relationships among them, that are presumed to exist in some area of interest. Actual information sources are linked separately with the terms of the ontology, using a definition language as well. CGI-scripts in Perl are computed to dynamically generate the HTML pages returned to the user for browsing. While this approach adds a certain level of automated linking to aid navigation, the DHE provides a generally larger set of links based solely on the database structure and entity-relationship schemas, as well as metadata. The DHE, however, does not provide customized templates for domain-specific navigational contexts. The DHE might integrate well with Geldof's approach to provide an additional level of functionality.

Moreover, many products have been released recently that aim to interface RDBMS and Web servers (Frank 1995). The solutions employed in these products require huge programming effort in SQL or a scripting language. The ease of integration with the DHE depends on how easy it is to parse application displays to identify the elements of interest, and to specify the commands to return to the application in the mapping rules. If the application has an API or marks the elements in the displays (as should become the custom as XML becomes more prevalent), building the application wrapper should be relatively easy.

The approaches presented above presuppose the hypertext designer's insight into the intrinsic semantics of the relational structure. A different approach was proposed by Papadopoulos et al (1996). Instead of relying on the relational schema of the database, a more semantically enriched Extended Entity-Relationship (EER) schema is semi-automatically produced, by incorporating a reverse engineering methodology. Hypertext views, consisting of node and link types, can be defined over the EER schema, while the SQL queries to instantiate them at runtime where automatically created, based on mapping information gathered during the reverse engineering process. Currently the DHE requires people to enter the entity-relationship schemas manually to the Database Schema Mapper Module (see 3.5). This application could
help to automate this process, and perhaps provide additional relationships, which the DHE could display for database applications.

While the hypermedia paradigm embodies an approach to structure and navigate information, it has several shortcomings. In particular, few hypermedia systems have focused on methodologies for information storage and retrieval. Database systems, on the contrary, are only concerned with storage and retrieval of information based on a formal model. They exhibit powerful methodologies for information storage, and effective indexing and querying. Furthermore they provide facilities such as transaction management, concurrency and access control as well as locking mechanisms.

The DHE should be able to streamline a company's software development efforts in several ways. It automatically supplements the organization's applications with hypermedia links, structuring, navigation and annotative functionality, and metadata. It also implements inter-application linking, as the university department example shows. Mapping rules can point to any accessible application. DHE also can speed the development of applications. Developers can offload link management, navigational structures (such as guided tours), user preference management and other features to DHE.
CHAPTER 2

THE DYNAMIC HYPERMEDIA ENGINE

A Web-based prototype of the Dynamic Hypermedia Engine (DHE) has been developed, which redesigns an older PC-based prototype (Bieber 1999). Figure 1 shows a screenshot of a database query result in the main frame. The DHE has added anchors to all parts of the query result, including the field names at the column heads. The user has clicked on “Counseling Center - department,” resulting in metadata for the element in the bottom center frame and a list of links in the bottom right-hand frame. Selecting any link will generate an SQL query to create the appropriate result. Currently the list of links includes only database structural links, such as finding the primary keys for this element. The bottom left-hand frame contains menus for any integrated application or DHE internal module. Links represent relationships and relationships have “meta-information” as well. Selecting the “¶” next to any link will provide metadata and a list of links for it. The DHE’s next release will provide these for menu items as well. The metadata frame currently displays the full Resource Description Framework (RDF) record. The DHE’s next release will preformat the metadata for display. Future versions will also filter and rank the links and metadata based on the user task and preferences.
As this demonstrates, link generation in the DHE does not result from any type of lexical analysis. Our focus is not on the display content of the link anchor, but rather on the application elements underlying each anchor. A “mapping rule” encodes each relationship found between two elements of interest at the “class level”. For example, suppose an application display shows the name of a university department. Departments generally have professors and courses taught (based on the standard entity-relationship diagram within a database system), as well as a Web page, an annual budget (within the accounting system), hires-in-progress (within the personnel system), a location on a map (within a geographic information system), etc. Individual mapping rules contain an algorithm or computation (set of commands) leading to the appropriate component in these respective systems. When the user selects a particular department, the DHE constructs these commands with the actual department instance selected and sends them to the appropriate destination system, which then retrieves— or more often generates— the resulting page. For example, one mapping rule could state that an element of type “department” would be related to an element of type “annual budget” through a relationship with the semantic type “annual budget for” and with a
parameterized command to retrieve annual budgets from the accounting system. Developers may take advantage of this to integrate database applications with other applications without altering their contents; they only have to add new mapping rules for the relevant element types.

The DHE executes concurrently with database management systems, database applications, and other applications such as the accounting system, providing automated link generation and other hypermedia functionality without altering them. Developers write an independent application "wrapper" and a set of mapping rules for each. Note that once a wrapper is written and the mapping rules are specified for each type of application (geographic information system, relational database management system, accounting package, etc.), the DHE will support all instances of that application in the future (new maps, database contents, budget sheets, etc.).

The DHE executes as follows. Applications or their wrappers connect to DHE through World Wide Web components, such as Servlets and JavaServer pages. It intercepts all messages passing between the application and its user interface, and uses the mapping rules to map each appropriate element of the message to a hypermedia anchor. The DHE's Web browser wrapper merges these anchors into the document being displayed and passes the resulting HTML document through the Web component servlet to the user's Web browser. When the user selects an anchor, the browser wrapper passes it to DHE, which returns a list of possible links (one for each appropriate relationship as determined by the mapping rules) and metadata. If the user selects a DHE link (e.g., to add an annotation or stage in a guided tour), the DHE processes it entirely. If the user selects a relationship with a destination in a known application, the DHE infers and instantiates the appropriate SQL queries or other application commands from the relationship's mapping rule and passes them to the target application for processing. If the user selects a user-created annotation or tour, etc., the DHE retrieves it. Thus the DHE automatically provides all hypermedia linking (as well as navigation) to applications, which remain hypermedia-unaware and in fact often entirely unchanged.

Figure 2 shows the DHE's logical engine architecture. Some major components are described here. The others are described on our project Web site (http://space.njit.edu:8001). The current prototype has been developed in Java. XML is the messaging format used for intra-
module communication, and RMI is the underlying protocol used to transfer these messages. While the browser wrapper currently produces HTML documents for display, a future version will produce XML documents, which take advantage of the Web's new XLink, and XPointer standards to handle anchors and links.

![DHE Architecture Diagram](image)

**Figure 2: DHE Architecture**

User Interface Wrappers serve three important functions: First, they translate the DHE's internal messages from DHE's standard format to a format the browser (or other User Interface or UI) can process, and vice versa. Second, they handle communication between the engine and the UI. Third, they implement any functionality DHE requires from the UI (e.g., maintaining parameters), which the UI cannot provide itself.

The Message Manager Module enables the communication between all DHE modules, routing all DHE internal messages.

The Mapping Rules Module maps the application data and relationships to hypermedia objects at run-time. The Mapping Rules Module maps the element instances in the virtual document to global element types (classes), and infers all relevant relationships (links) and metadata for the given element classes. These links and metadata are passed in messages to the UI Wrapper for display.
**Application Wrappers**, like user interface wrappers, manage the communication between DHE and their application systems, such as database applications and DBMS. They translate user requests from DHE’s internal format to the application’s programming interface (if any). They receive output from the application, convert it to the DHE format, mark the elements for the mapping rules module, and send it to DHE for eventual display on the UI.

**Other Hypermedia Functionality:** A series of other service modules will be implemented in future versions. Most will implement various kinds of hypermedia structuring, navigation and annotation functionality (Bieber et al. 1997; Conklin 1987; Nielsen 1995). Hypermedia structuring functionality includes local and global information overviews; node, link and anchor typing; as well as keywords, attributes and metadata on all of these. Navigation functionality includes structure-based query, sophisticated history-based navigation and bidirectional linking. Annotation functionality includes adding user-declared links, comments and bookmarks to dynamically-generated documents and displays.
CHAPTER 3
HYPERMEDIA SUPPORT FOR RELATIONAL DATABASE MANAGEMENT SYSTEMS

3.1 Introduction to the Relational Database Wrapper Module

One of the first tasks in providing hypermedia support is to intercept messages between the computational and user interface (UI) portions of the DMIS.

In the case of a Relational Database Management System (RDBMS), it provides only the computational portion. The UI portion is the responsibility of the database application. The RDBMS provides a standard way of requesting computational services (i.e. store, retrieve and analyze data) by means of Structured Query Language (SQL) statements. Database applications send SQL statements to an RDBMS. The RDBMS then executes these statements and returns the results of these statements back to the application, which then displays these results in a UI intuitive to its domain.

We have developed a service module, the Relational Database Wrapper Module (RDWM) that accepts requests to, execute SQL statements on the underlying database, and retrieve metadata of an element. It will also provide a UI allowing users to execute SQL statements and view results, metadata and all the relationships of the data affected by the SQL statement. This UI provides a hypertext-enriched view of the data stored in the database and also acts as the debug or test interface for our module.

Once a document containing the results of the original request has been created, our module identifies and marks all “elements of interest”. An “element of interest” is an entity that may have a relationship with another entity (element) or may have metadata of its own.

This document is then sent to the DHE Message Manager, which routes it to an intermediate module called the Mapping Rules Module (MRM). The MRM maintains a list of Mapping Rules, which are representations of the various inter-relationships between elements in a domain. It retrieves all the Mapping Rules that each marked-up “element of interest” may have and creates hyperlinks corresponding to each one of them.
The modified XML document containing the results of the original SQL statement (or the requested metadata) and hyperlinks is returned to the gateway, which then sends it onward to the database application. The database application thus receives hypertext-enriched result of a SQL statement from the RDBMS, and is free to interpret it in a manner suitable to its domain.

For example when the user uses the UI to execute a SQL statement a message containing the SQL statement is sent from the UIW to the RDWM (by way of the Message Manager). The RDWM executes the SQL statement on the underlying database, identifies the "elements of interest", marks them up and sends back a response containing this to the module that had sent the original request (i.e. the UIW). The Message Manager routes this message to the Mapping Rules Module, which applies the appropriate Mapping Rules and forwards the response to the UIW (as before, by way of the Message Manager).
The Message Flow is as:

![Diagram showing the message flow](image)

3.2 Elements of Interest in Database Systems.

3.2.1 Element Types.

The following types of elements exist in the RDBMS context:

- Columns.
- Tables.
- Indices.
- Stored Procedures.
- Catalogs.
- Schema.
- Drivers.
- Users.
- User Rights.
- Table and Column Privileges.
- JDBC Types.
- The RDBMS Product itself.

Any instance of the above types can be uniquely identified, have metadata, and have one or more relationships associated with it. Because any user may be interested in exploring that type of object in terms of its metadata or relationships, the RDWM must mark each of its instances as an “element of interest”.

3.2.2 Marking an Element of Interest.
Marking an item involves providing its unique identifier and its element type. When the results of the SQL statement is returned from the RDBMS, the RDWM must parse through this document and locate all elements and mark all instances with locator tags. An element’s locator tag references its unique identifier and its type.

Later on as the Message containing this document makes its way to the Mapping Rules Module it will use the element type information to find relationships and metadata for that element. If an element has at least one relationship or piece of metadata, then the Mapping Rules Module will specify that the UI Wrapper make it into a hyperlink.
3.2.3 Element Identifiers.

When the user follows a hyperlink, the action is passed on to the underlying DMISW (in our case the RDWM), as mentioned previously this action may either be a request for the hyperlinked element’s metadata or the user may be trying to follow a relationship with another element in which case it would be another SQL statement that has to be executed. In either case the element must be uniquely identified.

Like other DHE modules the RDWM uses the Uniform Resource Identifier (URI) syntax (Berners-Lee 1998) to define its element identification scheme. The generic URI syntax specifies that the name of the scheme must be specified followed by a colon (which acts as the delimiter), and followed by the scheme-specific part.

```scheme name>:\<scheme-specific-part>```

The DHE URI scheme is named “dhyme” (Dynamic Hypermedia Engine), and all URI’s must also contain the module name which identifies the domain the element belongs to.

For elements belonging to the RDWM domain we use the following syntax.

```
dhyme:rdwm:<element type>:<Database JDBC URL>:<element specific part>
```

The JDBC (Java Database Connectivity) URL of the database provides a way of identifying it so that the appropriate database driver will recognize it and establish a connection with it. (White 1999)

Thus a column called “DEPT” in the table “DEPARTMENT” in the database with the JDBC URL “jdbc:oracle:thin:@logic:1521:logic40” would have the following URI.

```
```

Because of the case insensitive nature of SQL statements this scheme too, is case insensitive.

3.2.4 Marked up Message.

The result of the SQL statement “select * from department” on the database the JDBC URL “jdbc:oracle:thin:@logic:1521:logic40” is as.
<?xml version="1.0"?>
<Doc>
  <DocID>rhyme:rdwm:QueryResults:select * from DEPARTMENT</DocID>
  <Type>rhyme:rdwm:QueryResult</Type>
</Doc>

<! [CDATA[
<table border="0">
<tr><td><b><#id1>DEPT</b></td>/<td><b><#id2>DESCRIPTION</b></td>/<td><b><#id2>DEPT</b></td>/</tr>
<tr><td colspan="2"><hr/></tr>
<tr><td><#id3>Computing Services</td>/</tr>
<tr><td><#id4>Counseling Center</td>/</tr>
<tr><td><#id5>Counseling Services</td>/</tr>
</table>]]>
</OutDoc>

</Outline>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>

</SubFrame Name="MAIN">
</Doc>
</MsgBody>
</MsgNo>RDWM 960332324364</MsgNo>
</Destination>UIW</Destination>
</RequestID>R60332321740</RequestID>
</MsgType>DISPLAYDOCUMENT</MsgType>
</Origin>RDWM</Origin>
</Current>RDWM</Current>
</GuidNo>RDWM 960332324364</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
</GuidNo>
3.3 Mapping Rules and Metadata for Relational Databases.

DHE specifies relationships based on the element type. The twelve types of elements noted above are interrelated. These inter-relationships are depicted below.

Each of these relationships are reflexive i.e., if a column has a “in table” relationship with a table then the table has a relationship called “has columns” with the column.
A mapping rule for each of the above relationships may be specified. The DHE (specifically the Mapping Rules Module) would then use the mapping rule to generate a link for each of the instances of the participating elements.

For example all instances of columns would have a hyperlink corresponding to the “in table” relationship that it has with a table. This link would contain the appropriate SQL command to generate the contents of the mapping rule’s endpoint, in this case the table that this column belongs to. Wan and Bieber (Wan 1996)(Bieber 1997) have given several mapping rules (called “bridge laws”) for relational databases.

3.3.1 Mapping Rules for Relational Database Management Systems.
Wan and Bieber (Wan 1996) have identified the following classes of Mapping Rules for databases:

Object Mapping Rules: These map the contents of database objects, as well as schemata and ER diagrams.

Structure Mapping Rules: These map database objects to their structural containers, for example a mapping rule for columns in a table, tables in a catalog etc. These would correspond to the “In Object” relationships mentioned in Figure # 5.

Operation Mapping Rules: These map SQL queries. Frequently used specific queries maybe mapped, and users who follow link generated from this mapping rule would retrieve a document that contained the results of the SQL statement being executed.

Schema-based Mapping Rules: These mapping rules map relationships between entities as defined in the database schema.

Meta-information Mapping Rules: Meta-information Mapping Rules define reference links to database object statistics such as field types, size, table size, referential constraints etc.

3.3.1.1 Mapping Rules examples.
The following Mapping Rules have been implemented in the current DHE prototype.

- Structure Mapping Rules:
• Mapping Rule\textsubscript{getTable} maps tables to a set of records.

• Mapping Rule\textsubscript{getTuple} maps the tuple a database value belongs to.

• Schema-based Mapping Rules:

  • Mapping Rule\textsubscript{getPrimaryKeys} maps the primary keys of a table.

  • Mapping Rule\textsubscript{getForeignKeys} maps the foreign keys in a table.

• Meta-information Mapping Rule:

  • MappingRule\textsubscript{getMetaInfo} maps the meta-information of a Database Object.

3.3.2 Metadata for Relational Database Management Systems.

Metadata is data about data. The key purpose of metadata is to facilitate and improve the retrieval of information (Miller 1998). Other uses of Metadata include providing semantic information about the data, encoding information (i.e. how to interpret the data), relationships with other resources etc. (Waugh 1997).

In a relational database, metadata is the representation of the objects defined in that database - specifically, the definitions of its tables, columns and business rules and transformational algorithms implemented as stored procedures and triggers (Gardner 1998).

The RDWM retrieves this metadata on demand. It is passed the URI of the element whose metadata is being asked for. It then uses the Java Database Connectivity (JDBC) API to retrieve information relevant to that element. The RDWM uses the Resource Description Framework (RDF) to model the metadata, and the RDF/XML serialization syntax to format it for transfer to other modules.

3.3.2.1 The Resource Description Framework.

RDF is a framework for describing and interchanging metadata it defines metadata in terms of "Resources", where each resource is any entity that can be uniquely identified, i.e., has a URI, and is defined by its properties (Lassila 1999).
Hence all “elements of interest” are resources and have associated metadata. Resources and elements are thus interchangeable in the DHE context and the element identifier corresponds to its URI. Thus there exists a Resource for every element type mentioned in section 3.2.1.

RDF also defines a “Property Type” as a resource that has a name and can be used to define a resource property. Consider the statement “The student’s name is John Doe”. “Student” is the resource, “name” is the property type and “John Doe” is the property.

3.3.2.1.1 RDF Schema for Databases.
However the semantics of the statement in the previous section is meaningful to the reader if and only if (s)he knows what the property type “name” actually means. RDF provides a schema mechanism, which allows authors (i.e. those who define what the resources and property types are) to define the terms that will be used in RDF statements and to give specific meaning to them. The RDF schema may thus be considered to be the dictionary to the RDF vocabulary in use (Brickley 2000).

To represent RDBMS resources in RDF, an RDF schema containing the various property types used by the RDWM has been defined in Appendix A.

3.3.2.1.2 RDF/XML Serialization Syntax
A formal syntax representing this metadata model is required to store instances of this model into machine-readable files and to communicate these instances among applications. XML is this syntax. RDF imposes formal structure on XML to support the consistent representation of semantics (Miller 1998).
For example consider the following model that represents metadata about a column called "Description" in the table "Department" on the Oracle instance on "logic".

The above model in RDF/XML serialization syntax is:

```xml
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <rdf:Description about="dhyme:rdwm:column:jdbc:oracle:thin:0logic:1521:logic40:Department:Description">
    <ORDINAL_POSITION>2</ORDINAL_POSITION>
    <CHAR_OCTET_LENGTH>500</CHAR_OCTET_LENGTH>
    <SQL_DATETIME_SUB>0</SQL_DATETIME_SUB>
    <TABLE_SCHEMA>Mohan</TABLE_SCHEMA>
    <DECIMAL_DIGITS></DECIMAL_DIGITS>
    <COLUMN_NAME>Description</COLUMN_NAME>
    <NUM_PREC_RADIX>10</NUM_PREC_RADIX>
    <COLUMN_SIZE>500</COLUMN_SIZE>
    <TABLE_CAT></TABLE_CAT>
    <COLUMN_DEF></COLUMN_DEF>
    <REMARKS></REMARKS>
    <IS_NULLABLE>YES</IS_NULLABLE>
    <DATA_TYPE>12</DATA_TYPE>
    <BUFFER_LENGTH>0</BUFFER_LENGTH>
    <SQL_DATA_TYPE>0</SQL_DATA_TYPE>
    <TABLE_NAME>Department</TABLE_NAME>
    <TYPE_NAME>VARCHAR2</TYPE_NAME>
  </rdf:Description>
</rdf:RDF>
```

Figure 6: Metadata of a Column
3.3.2.2 Architecture and Design of the Metadata Retrieval Subsystem

The Metadata Retrieval Subsystem is composed of two broad components.

- A set of utility classes responsible for storing Metadata and representing them in the RDF/XML serialization syntax, these classes comprise the “RDF Package”. These classes may be used outside the RDWM and indeed outside the DHE.

- A set of classes extending the functionality provided in the RDF package for use in the RDBMS context.

3.3.2.2.1 The RDF Package

The RDF Package essentially consists of two core classes the “Resource” class and the “ResourceProperty” class. These classes are present in the Java package “dhyme.metadata.rdf”.

</rdf:RDF>

Figure 7: RDF/XML Syntax
3.3.2.2.1 Object Oriented Analysis and Design

Figure 8: Class diagram of the "dhyme.metadata.rdf" package.
3.3.2.2.1.2 Features

- Each Resource Object has 0..n ResourceProperty Objects.

- Each ResourceProperty defines an attribute of the Resource, in turn a Resource Object is defined by the ResourceProperty Objects it "has".

- A Resource must have a Uniform Resource Identifier (URI). A URI uniquely identifies the Resource and by extension the ResourceProperty Objects it possesses.

- A ResourceProperty may be another Resource or a collection (referred to in RDF syntax as a "container") of Resource Objects. This collection may be a set of unordered Resource Objects (referred to in RDF syntax as a "bag") or it may be a set of ordered Resource Objects (referred to in RDF syntax as a "sequence"), or it may be a set of alternates, i.e. each Resource Object may substitute the other (referred to in RDF syntax as a "alternate" or "alt").

- A Resource Object has the ability to add, retrieve and remove ResourceProperty Objects.

- A Resource Object has the ability to express itself in the complete RDF-XML "serialization" syntax.

- A Resource Object has the ability to express itself in HTML and in Plain Text as an aid to debugging and for viewing metadata on browsers.

- ResourceWriter is a utility class used to "write" a collection of Resources into an Engine Message or express them in the RDF/XML serialization syntax.

- ResourceReader is a utility class that may be used to parse RDF resources from an Engine Message or from an RDF Element (i.e. the tag named rdf:RDF).
3.3.2.2.1.3 Recommended Usage
Create an instance of a Resource object and add/remove/change its properties by adding/removing ResourceProperty Objects to its collection of ResourceProperty Objects (as maintained in an internal Hashtable). The ability to serialize a Resource into its proper RDF/XML format is available by means of calls to rdfForm() and rdfForm(schemaName) on its instance.

Another approach is to subclass Resource class and then "load" ResourceProperty Objects in the constructor of the subclass. An example of this Usage is provided by the RDWM metadata retrieval subsystem.

3.3.2.2 The RDWM Metadata Retrieval Subsystem
This subsystem is an extension of the RDF package applied to the Relational Database Domain. The classes that comprise this subsystem are present in the "dhyme.metadata.database" package.
3.3.2.2.2.1 OBJECT ORIENTED ANALYSIS AND DESIGN

The core of this package is an abstract class called `DatabaseResource`, which in turn extends the `Resource` class. `DatabaseResource` provides an abstract method called `loadProperties`, which takes as an argument a `java.sql.DatabaseMetadata Object`.
• It is then the responsibility of each subclass of DatabaseResource to extract whatever metadata is relevant to it from the DatabaseMetadata instance passed to it.

• For example the ColumnResource class is responsible for extracting all metadata related to Columns in the Database, the TableResource class is responsible for extracting all metadata related to Tables in the Database.

3.3.2.2.2.3 RECOMMENDED USAGE
Instantiate the appropriate subclass of DatabaseResource by passing in an instance of a DatabaseMetaData object from the database that contains the Resource (Table, Column etc.), and use the inherited methods of the Resource class to obtain the metadata in RDF/XML serialized form.

3.4 The Relational Database Wrapper Module
3.4.1 Features.
The Relational Database Wrapper Module passes through the following states:
Figure 10: RDWM State Diagram
3.4.1.1 Sending Startup Messages.
Because of the DHE’s distributed nature, all modules must register themselves with the DHE Message Manager at startup time, so that the Message Manager knows which modules are currently active and able to receive and process requests.

3.4.1.1.1 Inputs.

3.4.1.1.1.1 PARAMETERS
- The Module-ID. This identifies the module and hence must be unique for every instance of this (and all other) module(s). The value of this is dependent on the Message Manager’s mechanism of registration. The Message Manager assigns a unique Module ID whenever any module tries to register itself. This module then proceeds to use this value as the remote object identifier when it binds itself to the local RMI Registry.
- IP address. Identifies the machine the RDWM is running on. This is required so that the Message Manager is able to perform an RMI “lookup” on this module.

3.4.1.1.2 Processing
Like all DHE modules the RDWM extends the Module class in the dhyme.gateway.module package. It thus inherits the startMe method. An invocation of this method with the URL of the Message Manager module sends the required Startup Messages.

The startMe method automatically discovers the IP address of the RDWM and sends this information to the Message Manager.

3.4.1.1.3 Outputs
Registration is essential for any DHE module to send and receive messages. Thus if registration fails, the RDWM exits with a message informing the user that registration has failed.
3.4.1.2 Receiving Request Messages

A module waits until it receives a request to perform a service. The services the RDWM provides are:

- Interact with the RDBMS.
- Retrieve metadata for an element
- Generate a user input form to enter SQL statements.

Like all DHE modules the RDWM receives requests by an invocation of `rcvMsg` a remote method it inherits from the `Module` class. The argument to this method is the Request Message itself.

The `rcvMsg` method goes on to call the method `processMsg`, which is an abstract method in the `Module` class. Like all DHE modules the RDWM provides an implementation of this method.

3.4.1.2.1 Inputs

An XML Message. This XML Message must contain the following two parameters:

- The source of the request, which will be used to determine the destination of the response of the RDWM.
- The action to be performed by the RDWM.

A sample Message:

```xml
<..prolog..>
<EngMsg>
  <RequestID>967145925626</RequestID>
  <MsgType>FOLLOWLINK</MsgType>
  <Origin:UIW</Origin>
  <Current:BLEM</Current>
  <MsgNo:UIW:967145925675</MsgNo>
  <Destination:RDWM</Destination>
  <User_name>guest:guest:Guest User</User_name>
  <MsgBody>
    <Subject>select</Subject>
    ........ other elements
```
The source of the message is defined in the element Origin. The action to be performed by the RDWM is defined in the element Subject.

The values of Subject is a command to execute a SQL statement, and will have one of the following values:

Select  Execute a select statement

Insert  Execute an insert statement.

Update  Execute an update statement.

Delete  Execute a delete statement.

Alter   DDL (Data Definition Language) command

Create  DDL command

A sample message to execute a SQL Select statement.

...... other elements
<MsgBody>
  <Subject>select</Subject>
  <BLid></BLid>
  <Parameter>
    <ParamName>Submit</ParamName>
    <ParamValue>Execute Query</ParamValue>
  </Parameter>
  <Parameter>
    <ParamName>query</ParamName>
</MsgBody>
<ParamValue>select * from publishers</ParamValue>
</Parameter>
</MsgBody>

All the above values of Subject require that a Parameter tag and a ParamName tag having the value query exist, and that it have an associated non-null ParamValue tag; which is the SQL statement that will be executed by the RDWM.

The value of the Subject tag may also be a command to retrieve the Metadata associated with a given Database Element (As defined in Section 3.3.2)

getMetaInfo Gets the metadata associated with a given RDBMS element.

A sample getMetaInfo Message:

...... other elements
<MsgBody>
  <Subject>getMetaInfo</Subject>
  <BLid></BLid>
  <Parameter>
    <ParamName>Type</ParamName>
    <ParamValue>dhyme:rdwm:value</ParamValue>
  </Parameter>
  <Parameter>
    <ParamName>Element_ID</ParamName>
  </Parameter>
</MsgBody>

For a getMetaInfo message, two parameter tags are required.

- One Parameter tag must have a ParamName with the value Type and an associated non-null ParamValue, which contains the type of the element whose metadata is being requested.

- The other Parameter tag must have a ParamName with the value Element_ID and an associated non-null ParamValue that contains the URI of the Element whose metadata is being requested.
Two other values of \textit{subject}, which do not perform any action on the underlying RDBMS, are also supported.

\begin{verbatim}
List                Lists all the values of \textit{subject} that are supported by the RDWM.
Display            Displays a User Interface to enter a Data Manipulation Statement.
\end{verbatim}

\subsection*{3.4.1.3 Extracting and Executing SQL statements}
Like all DHE modules the RDWM uses an XML parser to extract the action to be performed and SQL statement (if any) from the request message. If the request is to execute an SQL statement then the UI Wrapper will have embedded the actual SQL in a request message. (This is its default procedure for user input forms.) The RDWM then uses a persistent pool of JDBC (Java Database Connectivity) Connection Objects, to execute a SQL statement on, or retrieve metadata from the RDBMS.

\subsubsection*{3.4.1.3.1 Input}
An XML message containing the SQL statement to be executed. As mentioned in section 3.4.1.2.1 there are 6 cases the RDWM may receive a message containing a SQL statement to be executed, and in each case the SQL statement is in the \texttt{Parameter} tag of the message.

\subsubsection*{3.4.1.3.2 Processing}
The Java Programming Environment provides an API for performing RDBMS related tasks. This API, known as the Java Database Connectivity (JDBC) API, creates a programming-level interface for communicating with all relational databases in a uniform manner similar in concept to Microsoft's Open Database Connectivity (ODBC). The JDBC API is based on the X/Open SQL Call Level Interface, the same basis as that of ODBC. The core JDBC API consists of three interfaces: \texttt{Connection}, \texttt{Statement}, and \texttt{ResultSet}. A typical operation would establish a connection with a database get an instance of a \texttt{Connection},
then execute a query against that database using a `Statement` object, which would then return a `ResultSet`.

### 3.4.1.3.2.1 Connection Overhead

However the overhead time for establishing a database connection is typically around 1-3 seconds. This is the time it takes to locate the database server, establish a connection with it, and exchange login information. For applications where the database query times are large, this overhead is probably too small a fraction of the overall turn-around-time to be a critical issue. However when the application has to perform numerous short-term queries, database connection overhead can become a serious issue.

### 3.4.1.3.2.2 Connection Pool Oriented Architecture

The solution to this problem is to create a pool of persistent (reusable) database connections to be used by various application components as needed. This pool of connections is created and managed by a separate process or thread, the **connection broker**. An application component requests the broker to hand it a database connection. In addition, the broker manages the pool of connections, watches for locked or corrupted connections, logs events and performs other housekeeping tasks. Once the application component has completed its database request, the connection must be returned to the pool for reuse.

![Connection Pool Architecture](image)
3.4.1.3.2.3 CONNECTION POOL USAGE
The connection broker is in the class DBBroker in the package dhyme.utils.dbBroker. The connection broker is a singleton, i.e. only one instance of it ever exists in the life of the application/VM. When the application initializes, it must initialize the broker. Different application components then proceed to request the broker for a connection and once they are done with it they return this connection back to the broker.

- During Application Initialization:

  DBBroker.init(database parameters);

- When a Component needs a Connection:

  Connection conn = DBBroker.getInstance().getConnection();

  ....... .... do something with the connection

  DBBroker.getInstance().returnConnection(conn);

3.4.1.3.2.4 BROKER ARCHITECTURE
3.4.1.3.2.4.1 REQUESTING AND RETURNING CONNECTIONS
The Broker uses two Stacks to maintain the connections, the \textit{in} Stack represents those connections not yet handed out to an application component. The \textit{out} Stack represents those connections currently being used by various application components.

When a component requests a connection, and if one is not available (i.e. not in the \textit{in} Stack), instead of returning immediately (via a \texttt{java.sql.SQLException}), the broker waits a while and then looks again to see if a Connection is available. This process is repeated a number of times. The amount of time to wait for and the number of times this process will be repeated are parameters and are passed into the broker during initialization.

3.4.1.3.2.4.2 BROKER HOUSEKEEPING
It is possible that after the connection has been handed out to application components, it may have become corrupted during use and the application may not be able to return it back to the
pool. In such cases the broker should be able to remove that connection from the pool, and replace with a new uncorrupted connection.

To manage this housekeeping operation a `PoolManager` has been developed.

Deciding whether or not a connection is corrupted/locked is however not a trivial matter. A rudimentary way of deciding is to keep count of the length of time the connection has been in the `out` Stack. If this amount of time is extraordinarily “long” (i.e. greater than a max amount, passed in to the broker during initialization) and there are SQLWarnings for this connection, that would mean this connection is corrupted/locked, and hence should be closed immediately. A side effect of this is that the application component utilizing this connection
would get a `java.sql.SQLException` when it tries to execute queries or access a `ResultSet`, because the underlying resources would be closed.

### 3.4.1.3.2.4.3 OBJECT ORIENTED ARCHITECTURE

**Figure 14: Class Diagram of the `dhyme.utils.dbBroker.DBBroker` Class.**

#### 3.4.1.4 Identifying an Element of Interest

The Mapping Rules Module maps elements in the RDBMS domain into the hypertext environment. In order to perform this mapping, the Mapping Rules Module needs to identify which elements have any Mapping Rules or metadata associated with it. The RDWM (and all DMISWs) thus has to “mark” those elements.

As mentioned in section 3.2.1, there are 12 kinds of elements in the RDWM context and any instance of that is marked up.

#### 3.4.1.4.1 Input

The `ResultSet` from a query.
3.4.1.4.2 Processing
A ResultSetMarker class has been developed that iterates through the records returned in a ResultSet, formats those records into an HTML table for display and marks up each element with a locator tag. This HTML table is written into a DisplayDoc message for display in the UIW.

3.4.1.4.3 Output
The Response as an XML message with all the required elements marked up with locator tags, and the types and the URIs of these elements in the ListOfElements tag.

3.4.1.4.4 Format of Response
A DisplayDoc message.

3.4.1.5 Returning the Response
All DHE Modules extend the Module class and thus inherit the sendMsgToModule method. To send a message to any other DHE module, the source module must specify the ultimate destination of the module in the message, and call that method with the message as the argument.
3.4.2 Object Oriented Analysis

Figure 15: RDWM Class Diagram
3.4.3 Functionality.

3.4.3.1 Architecture.
As shown in the class diagram, the RDWM itself receives and sends messages, it parses the message and determines what command to execute. As mentioned in section 3.4.1.2.1, the command may either be to execute a query, to display a UI or to retrieve metadata. Another class of commands maybe issued as a result of a user following link created from a Mapping Rule, as detailed in section 3.3.1. These commands are also dealt with in the exact same manner.

To process this command the RDWM uses “handlers” devoted for that purpose, for example, a DisplayHandler is responsible for allowing a user to enter a SQL statement for execution.

The following sections detail the processing involved in each of these “handlers”, by a series of Collaboration diagrams that show the method calls in them. The arrows in each diagram signifies the direction of the method call, the arguments to that method (if any) is also labeled on each arrow.

3.4.3.2 Displaying the User Interface.
This is as a result of a display command being received by the RDWM.
3.4.3.3 Executing a select statement.
This is as a result of a `select` command being received by the RDWM.
3.4.3.4 Executing a non-select statement.

This is as a result of a delete, insert, update or DDL command having been issued from the UI.
3.4.3.5 Executing a Mapping Rule – Retrieve Primary Keys.

This is as a result of a user following through the “Get Primary Keys” Mapping Rule for a Column or Table.
3.4.3.6 Executing a Mapping Rule – Retrieve Foreign Keys.

This is as a result of a user following through the “Get Foreign Keys” Mapping Rule for a Column or Table.
3.4.3.7 Executing a Mapping Rule – Retrieve the Table a Column or a Value belongs to.

This is as a result of a user following through the “Get Table” Mapping Rule for a Column or Table.
3.4.3.8 Executing a Mapping Rule – Retrieve the Tuple a Value belongs to.

This is as a result of a user following through the “Get Tuple” Mapping Rule for a Database Value.
3.4.3.9 Retrieving Metadata.

This is as result of a user retrieving metadata for an element.
3.5 Enhanced Links through a Database Schema Wrapper

Most database applications provide no contextual information about the underlying schema of the database from which query results were retrieved. The DHE utilizes a dedicated Database Schema Mapper Module to add value to database applications by making this information explicit.
As shown above the three frames of the Database Schema Wrapper will subdivide the main frame of the DHE. The leftmost frame contains the database query results as before. The middle frame shows the relational (physical) schema behind the query results. The rightmost frame shows the original, non-normalized entity-relationship schema corresponding to the query results.

The DB Schema Mapper (DSM) runs in conjunction with the Relational Database Wrapper Module. As mentioned in section 3.3.2, the RDWM provides metadata by examining the physical schema and returns attributes such as names of columns, data types etc. To this metadata the DSM adds schematic information for the values retrieved from a database query.

The schematic information about a particular database has to be entered through a user interface by a system developer or administrator at the time the system is being designed or integrated with DHE. At runtime when a query has been issued, the DSM checks its internal database to see if the schematic information is available. If it is, then it allows the user the
ability to view the underlying E-R schema as well as the relational schema of the database from which the query result was retrieved.

The DSM receives messages either from the RDWM or from the User Interface Wrapper. The RDWM passes query result sets through the DSM to add schematic information. The DSM parses the original query to get the table name. It then queries its own internal database to see if it contains schema information for that table. If so, the DSM generates the E-R and relational schemas, marks whichever elements of interest each contains, and sends a message to the UI Wrapper to display these together with the regular query results.

The UI Wrapper sends the DSM messages when a user follows a DSM-related link. For example, when the user selects a table and chooses a link to highlight that table in the relational schema. The mapping rule corresponding to that link sends the appropriate command to the DSM. The DSM must follow its internal mapping from the RDWM URI scheme to the DSM URI scheme to identify the corresponding table element in the relational schema. Then the DSM creates a new display where it indicates that the UI Wrapper should highlight certain elements.

When parsing the RDWM's query results, the DSM marks the following as elements of interest in the corresponding E-R schema, and includes the properties shown as parameters:

1. E-R Database Schema
   - Name (of the database in the DSM internal database)

2. Entities.
   - Name.
   - Type (Weak Entity or Normal Entity).

3. Relationship.
   - Name
• Type (identifying relationship or weak entity relationship)

4. Attributes

• Name

• Type:
  1. Composite or Simple
  2. Multi-valued or Single-valued
  3. Stored or Derived
  4. Key or not

The mapping rules capture the inter-relationships amongst each of these elements. Given any particular database, entity, relationship or attribute in the E-R schema, the DSM will find the corresponding database, entities, relationships and attributes related to it in the E-R schema. Additional mapping rules will find the corresponding elements in the relational schema and the query result currently being displayed.

When parsing the RDWM's query results, the DSM marks the following as elements of interest in the corresponding relational schema, and includes the properties shown as parameters, with corresponding relationships:

1. Relational Database Schema

   • Name

2. Relation

   • Name

3. Attribute
• Primary Key

• Foreign key

4. Referential Integrity Constraint

• Relation name where the foreign key resides

• Relation name where the foreign key references

• Attribute name of the foreign key in the residing relation

  • Attribute name of the primary key of the referenced table.
CHAPTER 4

PROVIDING SUPPORT TO DATABASE APPLICATIONS

One of the roles Relational Database Wrapper Module (RDWM) as envisioned in the original architecture was to provide hypermedia enriched database services to existing as well as completely new Database Driven Applications.

4.1 Support to existing Database Applications

Developers can retrofit existing database applications to work with DHE. This section begins by describing such an integration. Then we describe the different kinds of links and metadata that DHE provides and conclude with a brief description of a system we have integrated with the DHE.

4.1.1 Motivation

Database applications that need to provide a hypertext-enriched view of data interact with the RDBMS via the RDWM. Such a view enables the user or application and database designers to view the underlying relationships that are not easily perceived, moreover by following links and retrieving metadata they gain an insight on the domain that the application serves.

4.1.2 Architecture

One of the first tasks in providing hypermedia support is to intercept messages between the computational and user interface (UI) portions of the application (Bieber and Kacmar 1995). In the case of a Relational Database Management System (RDBMS), the application comprises only a computational portion. The RDBMS provides a standard way of requesting computational services (i.e., store, retrieve and analyze data) by means of Structured Query Language (SQL) statements. The RDWM does this and provides its own interface for entering SQL queries and displaying their results.

Database applications build a customized interface and possibly a larger set of functionality around a RDBMS. Database applications, therefore, are responsible for their own UI. Database applications send SQL statements to their RDBMS. The RDBMS then executes these statements and returns the results of these statements back to the application, which then customizes and displays them.
If the UI displays are easy to parse, i.e., a developer could easily figure out which elements are in each display screen and how to pass back database and application commands to the application, then one could write an application wrapper that intercepts all displays and redirects them to DHE. This would satisfy our goal of providing automated hypermedia functionality with minimal change to the application. The DHE’s UI would then display the enhanced screens with hypermedia anchors as shown in Figure 1.

In general, Web-based database applications (Internet storefronts, catalogs etc.) have a clear distinction between their UI and the database; usually the UI uses a middleware to communicate with the database. The DHE may supplement or replace this middleware. However most legacy database applications (client server or mainframe based) usually don’t have such a clear distinction between their UI and the underlying database; most often the data and presentation is commingled. Integration with the DHE could be much more difficult in this case.

In the DHE architecture each database application would be a DMIS, which would have its corresponding DMIS Wrapper (DMISW). There are two integration points with an existing Database Application.

4.1.2.1 When the DMIS needs database services
This would be similar to supporting new database applications. The DMIS would route all database requests to the wrapper, which would pass it onto the rest of the Engine. Once the database request has been serviced, a hypermedia-enriched result of the database request would be provided to the DMIS, which could then use the UIW to display it. (See section 4.2 for more details)

4.1.2.2 When the DMIS interacts with its UI
All messages between the computational portion and the UI portion of the application would be intercepted by the Wrapper, which would route it to the engine to add hypermedia functionality to it and use the UIW to display the enriched UI (Chiu 1997; Bieber and Kacmar 1995). Assuming that a clear distinction exists between the computational and the interface portions of the application, and that the application allows these messages to be intercepted (Bieber and Kacmar 1995).
4.1.3 Case Study
This section describes the integration of a database application with the DHE.

4.1.3.1 Problem Description
The New Jersey Department of Transportation has an extensive database that contains commodity (coal, vegetable oil etc.) flow information between various counties in NJ and various zones in the Northeastern US. A rudimentary web interface to this database exists; we have created a DHE module that supplements this system and provides enhanced metadata, exposes the various interrelationships between the "elements of interest" in the system and provides additional functionality.

4.1.3.2 Architecture
The Freight Database Wrapper (FDW) is a DHE module that acts as the wrapper to the Freight system. Like other application wrappers described in previous sections, the FDW provides a gateway to application specific commands. Users may view the system through the DHE's User Interface and view reports on commodity flows between counties and zones, via a set of Menu Items and Mapping Rules.

The FDW may also use the RDWM to completely bypass the existing web based freight system and access the underlying database directly. This serves a twofold purpose:

- Metadata that is currently not provided by the system can be extracted directly from the database by means of SQL statements executed by the RDWM.

- New mapping rules maybe formulated, these mapping rules would correspond to SQL statements that would be executed on the freight database, thus providing additional functionality not available in the existing system.

The FDW may also be used in conjunction with the DSM, providing users with a view of the internal schema of the freight system's database.
4.2 Support to Applications being developed

This section describes how database applications could be developed quickly and easily if designed to take advantage of the DHE's infrastructure. A system that is being built using the DHE is also analyzed.

4.2.1 Motivation

In this role the DHE will provide access to a relational database for applications that need it. All requests to a database i.e., SQL statements that need to be executed on a database will be routed to the RDWM, which will execute the statement and return the hypermedia enriched results to the application.

This will allow any application developer to quickly develop a database driven application that is already enriched with hypermedia. The application developer will be able to use hypertext functionality (HTF) to visualize relationships between entities in the application’s domain, augment applications with, annotation and navigation functionality and other metadata that would normally be hidden inside the Relational Database. (Bieber and Vitali 1997)

4.2.1.1 Scalable Application

Most database applications are tied to a certain database schema, which must be known when the application is being designed. However by retrieving metadata and relationships (as a “list of links”) it is possible to create a generic database application that dynamically generates a UI corresponding to a database schema at runtime.

Thus instead of building a customized database application such as “Inventory” or “Purchase Order Tracking” where the UI and screen flow must be decided at the application design stage, it may be possible to build an application that utilizes the “list of links” to create a screen flow and metadata to generate the UI.

Because the application is not tied to a specific database schema, it is naturally scalable. Any time the Data Model or underlying “Business Rule” changes and a change has to be made to the schema, the application itself does not have to be modified, a change to an existing Mapping Rule or addition of a new one will suffice. This is in stark contrast to conventional application where a costly reengineering effort must be undertaken.
4.2.2 Architecture

To integrate with DHE, an application normally routes all database access requests to its application wrapper. In this case, because the application is being developed from the ground up, this wrapper maybe a part of the application itself. The wrapper (or wrapper portion of the application) would pass a DHE-formatted XML message to the RDWM to perform any database services requested by the application—usually the execution of a SQL statement or retrieval of metadata from the RDBMS being wrapped by the RDWM.

![Diagram](image.png)

**Figure 25: Providing Support to New Database Applications — Architecture**

RDWM would still be responsible for marking up any query results, passing the hypermedia-enriched document is then routed back to the application wrapper.
Figure 26: State Diagram

At this point the Application Wrapper may take the enriched document and translate it to the native application's native User Interface, and display it there. However, should the application developer decide to use DHE's user interface instead of writing its own, then the Application Wrapper could then add any additional content and pass the final display document to the Mapping Rules Module via the Message Manager. If it adds additional, non-
database elements, then the application wrapper should mark these up too, so they also may be made into anchors. Each of these additional elements types will require its own mapping rules for determining links and metadata. Many application commands can be moved into the link's mapping rules.

The application could also take advantage of other DHE services, such as the menu manager for displaying application specific menus and the user preference manager for managing users' sessions, login, profiles etc., not to mention the other hypermedia functionality that all applications receive.

4.2.3 Case Study

This section describes a proposal to develop a student paper review information system using the DHE.

4.2.3.1 Problem Description

As part of their coursework students must review a published paper, typically students would email the instructor the bibliographic reference to the article they would like to review and the instructor would approve or reject the request. If rejected the student would have to resubmit his/her request with a new article. If approved the student would review the article and email it to the instructor. The instructor would then grade the review and post it on the class message board. Because of the volume of email an instructor receives it becomes difficult for him/her to quickly approve or reject a request, because (s)he must go through previously approved requests and make sure that this article has not been approved for review by another student.

To ameliorate this situation we propose to develop a DHE module that would allow a student to request approval, check the status of the request, post a link to the review once completed and to check his/her grade, once the review is corrected. The instructor would be able to view pending requests, approve a request, view submitted reviews, assign grades, view reports on approved articles, submitted reviews etc. The instructor would also be able to perform administrative tasks such as adding students, deleting old reviews etc.
4.2.3.2 Architecture

The DHE only allows authenticated users access to the various modules (RDWM, etc.), if a user has not logged in, he/she is presented with a login screen, where a valid DHE username and password must be entered. At the start of semester, the instructor creates user accounts for all students. When a student logs in he/she presented with a menu, customized to the privileges that their user group has, in this case the student may have access only to this application and is not able to “play around” with other DHE modules.

Thus the paper review application module does not have to manage customized menus or user authentication issues; the DHE provides this service. Once a user has been authenticated any subsequent requests (i.e. view grades, approve a review request etc.) will always contain the username and the group the user belongs to. The application can thus perform or reject the action based on the privileges the user may have. The application developer may also register customized menus based on user groups; thus users belonging to the “student” group will not see menus for administrative tasks.

The DHE would also provide database access to this application. For example when the instructor wishes to view a report on all approved articles not yet submitted, he/she would click on the appropriate Menu Item and a message would be issued to the application module with a request to generate this report. The application module would then issue a message to the RDWM containing the appropriate SQL statement, the RDWM would execute the statement and return a document containing the results of the statement, with the “elements of interest” marked up. At this point the application module may simply forward this message to the UIW and it will be displayed in the instructor’s browser with hyperlinks in them. The instructor may follow this follow hyperlink to get more metadata for this element, and view a list of mapping rules. For example if the element is an article, the metadata may be the complete bibliographic reference to the article and a link to the article itself, and a mapping rule maybe a list of students who have requested this article

4.2.4 Advantages of using the DHE's infrastructure

The DHE thus speeds up the development process by freeing the developer from having to program user management, menu management, database access and above all hyperlink creation. All the developer would have to do is:
• Register a set of mapping rules corresponding to the additional functionality desired (i.e. view a list of students who have requested a specific article, view a list of pending approval requests for a given section etc.).

• Create a set of display screens.

• Formulate a set of SQL statements corresponding to each of the main commands (i.e. retrieve grades for a student, display submitted reviews etc.)

• Write a DHE Module that receives requests for each of the main commands and sends the appropriate SQL statement to the RDWM.
CHAPTER 5
USING THE DHE AS A DATA WAREHOUSING SOLUTION

This section proposes integrating a data warehousing application within the DHE infrastructure, one of our future research topics. This would provide hypermedia support to data warehousing applications, as well as facilitate linking among data warehouses and other applications.

5.1 Introduction

A data warehouse is a subject-oriented, integrated, time-variant and non-volatile collection of data in support of management decision-making (Inmon 1996). A data warehouse is a repository of information built using data from diverse, and often departmentally isolated, application systems within an organization so this data can be modeled and analyzed by managers (Johnson 1999; Inmon 1996). Data warehouses usually are customized for a particular enterprise. Most vendors offer platforms on which enterprise data warehouses (or smaller datamarts) may be built.

5.1.1 Data Warehouse Metadata

Data warehouse architectures integrate a metadata repository that contains:

- Administrative metadata: source databases and their contents; gateway description; warehouse schema, view and derived data definitions; dimensions and hierarchies; predefined queries and reports; data mart locations and contents; data partitions; data extraction, cleaning, transformation rules, default values; data refresh and purge rules; user profiles, user groups; security

- Business metadata: business terms and definitions; ownership of data; charging policies

- Operational metadata: data lineage (history of migrated data and sequence of transformation applied); currency of data: active, archived, purged; monitoring information: warehouse usage statistics, error reports, audit trails
5.1.2 Motivation
The DHE maybe used to offer a complete end-to-end solution with the added benefit of obtaining hypertext functionality without any additional effort.

Data warehouses are typically used for on-line analytical processing (OLAP). The key structure of a data warehouse always contains some element of time and some dimension hierarchies. OLAP queries are complex. They involve grouping and aggregation. A single OLAP query can lead to several closely related queries (Chaudhuri 1997). The visualization of an OLAP query result using DHE will involve links between data from one hierarchy level to the other and links SQL subqueries contained in the OLAP query. In addition, an OLAP query can result in a large collection of data with several dimensions. In the rest of this section we concentrate on the loader module.

5.2 Functionality
The data warehouse has two broad functions:

- Accessing the data from the data warehouse.

- Loading the data from the operational systems into the data warehouse.

Like all Data Warehouses a DHE based Data Warehouse will use a DBMS as the underlying store, thus functionality of accessing data from a data warehouse will be similar to any other database application and has been dealt in previous sections.

5.2.1 The Loader Module.
To load data into the data warehouse a loader module will have to be designed and developed. The Loader Module will be like any other DHE module, and will perform the following functions:

5.2.1.1 Mapping data from Operational Systems to the Data Warehouse
The operational systems store data in their own structure, encoding etc. this has to be mapped to the data warehouse’s format which is consistent across all operational systems.
5.2.1.2 Extracting Metadata from Operational Systems.
Metadata is the road map or blueprint to the data in the data warehouse, and needs to be operational. Also, metadata needs to be preserved for analysis once it has been loaded (Gardner 98). Metadata may include:

- The structure of the operational data.
- Relationships in the operational data.
- Other user-defined metadata.

5.2.1.3 Eliminating Noise
Operational data may contain data irrelevant to the warehouse (i.e., is noise), this data needs to be eliminated before loading.

5.2.1.4 Architecture
The Loader Module is supplied a template (an XSL stylesheet) that maps data from the operational systems' format to the data warehouse's format. It processes the template and maps data from the operational system to the data warehouse. Any data not specified in the template is noise and will be eliminated.

Once the extraction process is complete, the Loader Module sends a message to the RDWM containing the data to be loaded as well as the metadata. The RDWM then loads this into the data warehouse. Figure 27 describes the extraction process.

An argument could be made that a Loader Module is not required and the DHE is simply used to access data from the warehouse. However this approach would not allow the DHE to retrieve metadata from the operational systems. Moreover a complete end-to-end solution requires that we provide a Loader Module.
Figure 27: DHE Data Warehouse Loader Module
CHAPTER 6

FUTURE RESEARCH PLANS - INTEGRATING DATA MINING

6.1 Introduction
Currently DHE determines links from the mapping rules. Because the person who develops the application wrapper also writes the mapping rules at the same time, the types of relationships DHE finds are known ahead of time. Data mining brings the opportunity of a new kind of dynamic linking. Data mining searches large databases for relationships and global patterns and relationships that are not immediately obvious, such as a relationship between patient data and their medical diagnosis (Holsheimer 1994).

Data mining tools discover these relationships or models at runtime as opposed to design time. Thus, the DHE must request the Data Mining Tool to discover the relationships for an element of interest at runtime, and then use these discovered relationships to create hyperlinks.

Of course, in addition, the DHE could provide hypertext functionality to commercial Data Mining Tools. A commercial data mining tool would have a wrapper written for it, just as with any other application. Oracle Corporation’s “Darwin”™ is one of the most well known commercial Data Mining tools. It provides a comprehensive API (in C++) to access its data mining functions, it would be easy to write a wrapper that would be able to interface with it. (Tamayo 1997), and thus seems an ideal candidate to develop a DHE wrapper for.

6.2 Mapping Rules for Data Mining
As mentioned in previous sections the DHE uses logical rules called mapping rules to provide hypertext functionality to the components of the DMIS. Mapping Rules map the objects defined in the DMIS such as models, relationships etc. to objects in the hypermedia elements such as links etc. (Bieber 1995)

These Mapping Rules need to be defined when the DMIS Wrapper is being developed which means the models, relationships that are going to be mapped should be known. Normally this is not a problem since the relationships in the various elements/nodes/entities in a DMIS are well known and determined when the DMIS is being designed.
The problem with a Data Mining Tool is that it discovers these relationships or models at runtime as opposed to design time. Which means the Mapping Rules in this case cannot be defined and entered into the Engine when the Data Mining Wrapper is being developed.

One way of solving this would be define a generic set of Mapping Rules for Data Mining, which in turn would generate other Mapping Rules dynamically (Two step Mapping Rules).

6.2.1 Two Step Mapping Rules.

For each element in a document generated by the DMIS (i.e. the Data Mining Tool), the Mapping Rules module would ask it to generate Relationships for elements of the same type.

For example, consider that the Data Mining Tool is operating on a mass of data related to Automobile Tire Replacements that contain information on the customer who returned the tire(s), the reason for return, the type of vehicle etc. all accessible via the tire's serial number, and it operates on another mass of data relating the tire to the place of manufacture, the batch number and the composition of the components used. When asked to return possible relationships for a tire it would automatically return correlations between the point of manufacture, the vehicle the automobile tire was placed on and the composition of the components. The Mapping Rules Module would then generate Mapping Rules for each of these relationships and display them as hyperlinks allowing the user to quickly and navigate these relationships and obtain an overall view of the detailed data available to him/her.
APPENDIX A

RDF RESOURCE PROPERTIES USED BY THE RDWM

Legend:
- Resource - The type of the resource the property belongs to.
- Resource Property - The name of the property, also the tag used in the RDF/XML serialized metadata representation.
- Type - The type of the Property
- Range - Range of the property if any.
- JDBC Method - The JDBC API method used to retrieve this information
- RDF:Comments - Explanatory comments on the RDF/XML tag.
- JDBC Class - The JDBC API class the JDBC Method belongs to.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Resource Property</th>
<th>Type</th>
<th>Range</th>
<th>JDBC Method</th>
<th>RDF:Comments</th>
<th>JDBC Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catalog</td>
<td>NAME</td>
<td>String</td>
<td></td>
<td>getCatalog()</td>
<td>Name of Catalog</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>CHAR_OCTET_LENGTH</td>
<td>int</td>
<td></td>
<td>getColumns()</td>
<td>Maximum number of bytes in the column, valid for char types only.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>COLUMN_PRIVILEGE</td>
<td>Privilege</td>
<td></td>
<td>getColumnPrivileges()</td>
<td>Privileges for this Column</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>DECIMAL_DIGITS</td>
<td>int</td>
<td></td>
<td>getColumns()</td>
<td>Number of Fractional Digits</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>DEFAULT</td>
<td>String</td>
<td></td>
<td>getColumns()</td>
<td>Default Value for the column</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>IN_TABLE</td>
<td>Table</td>
<td></td>
<td>getColumns()</td>
<td>Which Table this Column belongs to.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>IS_NULLABLE</td>
<td>String</td>
<td>NO, YES, UNKNOWN</td>
<td>getColumns()</td>
<td>Whether or not a column can be null or not.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>IS_PSEUDO_COLUMN</td>
<td>String</td>
<td>Unknown, NotPseudo, Pseudo</td>
<td>getBestRowIdentifier()</td>
<td>Whether or not this is a pseudo column or not. (eg. Oracle ROWID)</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>JDBC_TYPE</td>
<td>JDBCType</td>
<td></td>
<td>getColumns()</td>
<td>JDBC data type.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>NAME</td>
<td>String</td>
<td></td>
<td>getColumns()</td>
<td>Column Name</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>ORDINAL_POSITION</td>
<td>int</td>
<td></td>
<td>getColumns()</td>
<td>Index of the column in table (first column is 1)</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>RADIX</td>
<td>int</td>
<td></td>
<td>getColumns()</td>
<td></td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>RDF:COMMENTS</td>
<td>RDF:Comments</td>
<td></td>
<td>getColumns()</td>
<td>Explanatory Comment on the Column</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>RDF:LABEL</td>
<td>RDF:Label</td>
<td></td>
<td>getColumnName()</td>
<td>Suggested Column Title for Column</td>
<td>ResultSetMetadata</td>
</tr>
<tr>
<td>Column</td>
<td>SIZE</td>
<td>int</td>
<td></td>
<td>getColumns()</td>
<td>Column Size. For char or date types this is the max number of characters; for numeric or decimal types this is the precision</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Driver</td>
<td>IDENTIFIER_QUOTE_STRING</td>
<td>String</td>
<td></td>
<td>getIdentifierQuoteString()</td>
<td>String used to quote SQL identifiers.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Driver</td>
<td>IS_NULL_PLUS_NON_NULL</td>
<td>boolean</td>
<td>true/false</td>
<td>nullPlusNonNullsNull()</td>
<td>Whether concatenation of a NULL value and a non-NULL value returns a NULL value.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Driver</td>
<td>MAX_CONNECTION</td>
<td>int</td>
<td></td>
<td>getMaxConnections()</td>
<td>Maximum number of connections that can be made to the database through this driver instance.</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Driver</td>
<td>NAME</td>
<td>String</td>
<td></td>
<td>getDriverName()</td>
<td>JDBC Driver Name</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Driver</td>
<td>VERSION</td>
<td>String</td>
<td></td>
<td>getDriverVersion()</td>
<td>JDBC Driver Version</td>
<td>DatabaseMetadata</td>
</tr>
<tr>
<td>Index</td>
<td>Column</td>
<td>Type</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------</td>
<td>----------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CARDINALITY</td>
<td>int</td>
<td>number of unique values in the index if TYPE is tableIndexStatistic then this is the number of rows in the table.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>FILTER_CONDITION</td>
<td>String</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>IN_TABLE</td>
<td>Table</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>INDEX_QUALIFIER</td>
<td>String</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>INDEX_TYPE</td>
<td>String</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>IS_NON_UNIQUE</td>
<td>boolean</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>NAME</td>
<td>String</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>ORDINAL_POSITION</td>
<td>int</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>PAGES</td>
<td>int</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index</td>
<td>SORT_SEQUENCE</td>
<td>String</td>
<td>get IndexInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAN_BE_AUTO_INCREMENT</td>
<td>boolean</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CAN_BE_MONEY_VALUE</td>
<td>boolean</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CASE_SENSITIVE</td>
<td>boolean</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CREATE_PARAMS</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DBMS_NAME</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>JDBC_NAME</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LITERAL_PREFIX</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LITERAL_SUFFIX</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LOCAL_NAME</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MAX_SCALE</td>
<td>int</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MIN_SCALE</td>
<td>int</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NULLABLE</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRECISION</td>
<td>int</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>RADIX</td>
<td>int</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SEARCHABLE</td>
<td>String</td>
<td>get TypeInfo()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Indicates whether it is possible to use a WHERE clause based on this type.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JDBC Type</td>
<td>Type</td>
<td>Description</td>
<td>Method</td>
<td>Database Metadata</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>-------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UNSIGNED</td>
<td>boolean</td>
<td>true/false</td>
<td>getTypeId()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Whether or not this type is unsigned or not.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAN GRANTEE_GRANT</td>
<td>boolean</td>
<td>true/false</td>
<td>getColumnPrivileges()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIVILEGE</td>
<td></td>
<td>Whether or not the grantee can grant access to others.</td>
<td>getTablePrivileges()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIVILEGE_GRANTEE</td>
<td>User</td>
<td>The User who is granted access.</td>
<td>getColumnPrivileges(), getTablePrivileges()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIVILEGE_GRANTOR</td>
<td>User</td>
<td>The User who grants access.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRIVILEGE_TYPE</td>
<td>String</td>
<td>Type of Access (SELECT, INSERT, UPDATE, REFERENCES etc.)</td>
<td>getColumnType(), getTablePrivileges()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>IN_CATALOG</td>
<td>Catalog</td>
<td>getProcedures()</td>
<td>Procedure Catalog</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>IN_SCHEMA</td>
<td>Schema</td>
<td>getProcedures()</td>
<td>Procedure Schema</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>NAME</td>
<td>String</td>
<td>getProcedures()</td>
<td>Name of the Procedure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td>TYPE</td>
<td>String</td>
<td>getProcedures()</td>
<td>Whether or not Procedure returns a Result</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>CATALOG_SEPARATOR</td>
<td>String</td>
<td>getCatalogSeparator()</td>
<td>Separator between catalog and Table names</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>CATALOG_TERM</td>
<td>String</td>
<td>getCatalogTerm()</td>
<td>Database vendor's term for Catalog Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>CATALOGS</td>
<td>Bag of Catalogs</td>
<td>getCatalogs()</td>
<td>All Catalog Names available in the database (Not in order)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>DEFAULT TRANSACTION ISOLATION</td>
<td>String</td>
<td>getDefaultTransactionIsolation()</td>
<td>Default Transaction Isolation Level. (By Isolation we mean atomicity of a Transaction)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>DRIVER</td>
<td>Driver</td>
<td>getDriverName()</td>
<td>JDBC Driver used to connect to the Database</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>DOES DDL IN TRANSACTION_COMMIT</td>
<td>boolean</td>
<td>dataDefinitionCommit(), dataDefinitionCausesCommit()</td>
<td>Does a DDL in a Transaction cause a Commit?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>EXTRA_NAME_CHARACTERS</td>
<td>String</td>
<td>getName(), getExtraNameCharacters()</td>
<td>ASCII Special Characters that can be used in names beyond a-z, 0-9 and _</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_BLOB_IN_MAX ROW SIZE</td>
<td>boolean</td>
<td>doesMaxRowsSizeIncludeBlobs()</td>
<td>Does the value returned by getMaxRowsSize include LONGVARCHAR and LONGVARBINARY blobs?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_CATALOG_AT_START</td>
<td>boolean</td>
<td>isCatalogAtStart()</td>
<td>Does the Catalog Name appear at the start of a fully qualified Table Name?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_DDL_IGNORED_IN_TRANSACTION</td>
<td>boolean</td>
<td>dataDefinitionIgnoredInTransactions()</td>
<td>Is a DDL ignored in a Transaction?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_NULL_SORTED_AT_END</td>
<td>boolean</td>
<td>nullsAreSortedAtEnd()</td>
<td>Are Null values sorted at the end regardless of sort order?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_NULL_SORTED_HIGH</td>
<td>boolean</td>
<td>nullsAreSortedHigh()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>IS_READ_ONLY</td>
<td>boolean</td>
<td>isReadOnly()</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>MAX_BINARY_LITERAL_LENGTH</td>
<td>int</td>
<td>getMaxBinaryLiteralLength()</td>
<td>How many hex characters can you have in an inline binary literal?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>MAX_CATALOG_LENGTH</td>
<td>int</td>
<td>getMaxCatalogLength()</td>
<td>What's the maximum length of a catalog name?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>MAX_CHAR_LITERAL_LENGTH</td>
<td>int</td>
<td>getMaxCharLiteralLength()</td>
<td>Maximum number of characters allowed in a character literal.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMN_NAME_LENGTH</td>
<td>int</td>
<td>getMaxColumnNameLength()</td>
<td>Maximum number of characters allowed in column Name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------</td>
<td>-----</td>
<td>---------------------------</td>
<td>----------------------------------------------------</td>
<td>------------------</td>
<td>------</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMNS_IN_GROUP_BY</td>
<td>int</td>
<td>getMaxColumnsInGroupBy()</td>
<td>Maximum number of columns allowed in GROUP BY clause.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMNS_IN_INDEX</td>
<td>int</td>
<td>getMaxColumnsInIndex()</td>
<td>Maximum number of columns allowed in an index.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMNS_IN_ORDER_BY</td>
<td>int</td>
<td>getMaxColumnsInOrderBy()</td>
<td>Maximum number of columns allowed in an ORDER BY clause.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMNS_IN_SELECT</td>
<td>int</td>
<td>getMaxColumnsInSelect()</td>
<td>Maximum number of columns allowed in a SELECT clause.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_COLUMNS_IN_TABLE</td>
<td>int</td>
<td>getMaxColumnsInTable()</td>
<td>Maximum number of columns allowed in a table.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_CURSOR_NAME_LENGTH</td>
<td>int</td>
<td>getMaxCursorNameLength()</td>
<td>Maximum number of characters that can be used in Cursor Name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_INDEX_LENGTH</td>
<td>int</td>
<td>getMaxIndexLength()</td>
<td>Maximum number of bytes allowed in an index.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_PROCEDURE_NAME_LENGTH</td>
<td>int</td>
<td>getMaxProcedureNameLength()</td>
<td>Maximum number of characters allowed in a procedure name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_ROW_SIZE</td>
<td>int</td>
<td>getMaxRowSize()</td>
<td>Maximum number of bytes allowed in a single row.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_SCHEMA_NAME_LENGTH</td>
<td>int</td>
<td>getMaxSchemaNameLength()</td>
<td>Maximum number of characters allowed in a schema name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_STATEMENT_LENGTH</td>
<td>int</td>
<td>getMaxStatementLength()</td>
<td>Maximum number of characters allowed in an SQL statement.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_STATEMENTS</td>
<td>int</td>
<td>getMaxStatements()</td>
<td>Maximum number of active statements to this database that may be open at the same time.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_TABLE_NAME_LENGTH</td>
<td>int</td>
<td>getMaxTableNameLength()</td>
<td>Maximum number of characters allowed in a table name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_TABLES_INSELECT</td>
<td>int</td>
<td>getMaxTablesInSelect()</td>
<td>Maximum number of Tables allowed in a SELECT clause.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>MAX_USER_NAME_LENGTH</td>
<td>int</td>
<td>getMaxUserNameLength()</td>
<td>Maximum number of characters allowed in a user name.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>NAME</td>
<td>String</td>
<td>getDatabaseProductName()</td>
<td>Product Name for this database.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>NUMERIC_FUNCTIONS</td>
<td>String</td>
<td>getNumericFunctions()</td>
<td>Comma Delimited List of Math Functions</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>PROCEDURE_TERM</td>
<td>String</td>
<td>getProcedureTerm()</td>
<td>Database vendor's preferred term for &quot;procedure&quot;</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>SCHEMA_TERM</td>
<td>String</td>
<td>getSchemaTerm()</td>
<td>Database vendor's preferred term for &quot;schema&quot;</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>SCHEMAS</td>
<td>Bag Of Schema</td>
<td>getSchemas()</td>
<td>Schema in the Database.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>SEARCH_STRING_ESCAPE_PATTERN</td>
<td>String</td>
<td>getSearchStringEscapePattern()</td>
<td>String that can be used to escape &quot;&quot;, &quot;&quot;, or &quot;%&quot; wildcards in the string search pattern used for catalog search parameters.</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>SQL_KEYWORDS</td>
<td>String</td>
<td>getSQLKeywords()</td>
<td>Comma-separated list of keywords used by the database that are not SQL-92 keywords</td>
<td>DatabaseMetaData</td>
<td>Data</td>
</tr>
<tr>
<td>Product</td>
<td>STORE LOWER_CASE_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesLowerCaseIdentifiers()</td>
<td>Does the database treat mixed case unquoted SQL identifiers as case insensitive and store them in lower case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------</td>
<td>---------</td>
<td>------------</td>
<td>-----------------------------</td>
<td>-------------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Product</td>
<td>STORE_LOWER_CASE_QUOTED_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesLowerCaseQuotedIdentifiers()</td>
<td>Does the database treat mixed case quoted SQL identifiers as case insensitive and store them in lower case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>STORE_MIXED_CASE_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesMixedCaseIdentifiers()</td>
<td>Does the database treat mixed case unquoted SQL identifiers as case insensitive and store them in lower case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>STORE_MIXED_CASE_QUOTED_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesMixedCaseQuotedIdentifiers()</td>
<td>Does the database treat mixed case quoted SQL identifiers as case insensitive and store them in lower case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>STORE_UPPER_CASE_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesUpperCaseIdentifiers()</td>
<td>Does the database treat mixed case unquoted SQL identifiers as case insensitive and store them in upper case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>STORE_UPPER_CASE_QUOTED_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>storesUpperCaseQuotedIdentifiers()</td>
<td>Does the database treat mixed case quoted SQL identifiers as case insensitive and store them in upper case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>STRING_FUNCTIONS</td>
<td>String</td>
<td></td>
<td>getStringFunctions()</td>
<td>Comma-separated list of string functions.</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT ALTER TABLE WITH ADD COLUMN</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsAlterTableWithAddColumn()</td>
<td>&quot;ALTER TABLE&quot; with add column supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT ALTER TABLE WITH DROP COLUMN</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsAlterTableWithDropColumn()</td>
<td>&quot;ALTER TABLE&quot; with drop column supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_ANSI_92_ENTRY_LEVEL_SQL</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsANSI92EntryLevelSQL()</td>
<td>Is the ANSI92 entry level SQL grammar supported? All JDBC Compliant drivers must return true.</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_ANSI_92_FULL_SQL</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsANSI92FullSQL()</td>
<td>Is the ANSI92 full SQL grammar supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_ANSI_92_INTEMEDIATE_SQL</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsANSI92IntermediateSQL()</td>
<td>Is the ANSI92 intermediate SQL grammar supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CATALOG_IN_DML</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCatalogInDataManipulation()</td>
<td>Can a catalog name be used in a data manipulation statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CATALOG_INDEX_DEFINITIONS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCatalogIndexDefinitions()</td>
<td>Can a catalog name be used in an index definition statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CATALOG_PRIVILEGE_DEFINITIONS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCatalogInPrivilegeDefinitions()</td>
<td>Can a catalog name be used in a privilege definition statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CATALOG_PROCEDURE_CALLS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCatalogProcedureCalls()</td>
<td>Can a catalog name be used in a procedure call statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CATALOG_TABLE_DEFINITIONS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCatalogTableDefinitions()</td>
<td>Can a catalog name be used in a table definition statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_COLUMN_ALIASING</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsColumnAliasSupport()</td>
<td>Is column aliasing supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CONVERT</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsConvert()</td>
<td>Is the CONVERT function between SQL types supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CORE_SQL</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCoreSQLGrammar()</td>
<td>Is the ODBC Core SQL grammar supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_CORRELATED_SUBQUERIES</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsCorrelatedSubqueries()</td>
<td>Are correlated subqueries supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT DDL_AND_DML_IN_TRANSACTION</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsDataDefinitionAndDataManipulationStatements()</td>
<td>Are both data definition and data manipulation statements within a transaction supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------</td>
<td>---------</td>
<td>-----------</td>
<td>------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_DIFFERENT_TABLE_CORRELATION_NAMES</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsDifferentTableCorrelationNames()</td>
<td>Are table correlation names supported, are they restricted to be different from the names of the tables?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_EXPRESSIONS_IN_ORDER_BY</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsExpressionInOrderBy()</td>
<td>Are expressions in &quot;ORDER BY&quot; lists supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_EXTENDED_SQL_GRAMMAR</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsExtendedSQLGrammar()</td>
<td>Is the ODBC Extended SQL grammar supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_FULL_OUTER_JOINS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsFullOuterJoin()</td>
<td>Are full nested outer joins supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_GROUP_BY</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsGroupBy()</td>
<td>Is some form of &quot;GROUP BY&quot; clause supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_GROUP_BY_BEYOND_SELECT</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsGroupByBeyondSelect()</td>
<td>Can a &quot;GROUP BY&quot; clause add columns not in the SELECT provided it specifies all the columns in the SELECT?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_GROUP_BY_UNRELATED</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsGroupByUnrelated()</td>
<td>Can a &quot;GROUP BY&quot; clause use columns not in the SELECT?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_INTEGRITY_ENHANCEMENT_FACILITY</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsIntegrityEnhancementFacility()</td>
<td>Does the SQL Integrity Enhancement Facility support?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT LIKE_ESCAPE_CLAUSE</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsLikeEscapeClause()</td>
<td>Is the escape character in &quot;LIKE&quot; clauses supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_LIMITED_OUTER_JOINS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsLimitedOuterJoins()</td>
<td>Is there limited support for outer joins?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_MINIMUM_SQL_GRAMMAR</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsMinimumSQLGrammar()</td>
<td>Is the ODBC Minimum SQL grammar supported? All JDBC: Compliant drivers must return true.</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_MIXED_CASE_IDENTITY</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsMixedCaseIdentity()</td>
<td>Does the database treat mixed case unquoted SQL identifiers as case sensitive and as a result store them in mixed case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_MIXED_CASE_QUOTED_IDENTIFIER</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsMixedCaseQuotedIdentifier()</td>
<td>Does the database treat mixed case quoted SQL identifiers as case sensitive and as a result store them in mixed case?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_MULTIPLE_RESULTSETS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsMultipleResultSets()</td>
<td>Are multiple ResultSets from a single execute statement supported?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_MULTIPLE_TRANSACTIONS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsMultipleTransactions()</td>
<td>Can we have multiple transactions open at once (on different connections)?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_NON_NULLABLE_COLUMNS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsNonNullColumns()</td>
<td>Can columns be defined as non-null?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_OPEN_Cursors ACROSS_COMMIT</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOpenCursorsAcrossCommit()</td>
<td>Can cursors be defined as open across commits?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_OPEN_Cursors ACROSS_ROLLBACK</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOpenCursorsAcrossRollback()</td>
<td>Can cursors remain open across rollbacks?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_OPEN_STATEMENTS ACROSS_COMMIT</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOpenStatementsAcrossCommit()</td>
<td>Can statements remain open across commits?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_OPEN_STATEMENTS_ACROSS_ROLLBACK</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOpenStatementsAcrossRollback()</td>
<td>Can statements remain open across rollbacks?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>SUPPORT_ORDER_BY_UNRELATED</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOrderByUnrelated()</td>
<td>Can an &quot;ORDER BY&quot; clause use columns not in the SELECT statement?</td>
<td>DatabaseMetaData</td>
</tr>
<tr>
<td>Product</td>
<td>Type</td>
<td>Value</td>
<td>Description</td>
<td>Supported?</td>
<td>DatabaseMetaData Type</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------</td>
<td>------------</td>
<td>-------</td>
<td>--------------------------------------------------</td>
<td>------------</td>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td>SUPPORT OUTER JOINS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsOuterJoins()</td>
<td>Is some form of outer join supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT POSITIONED DELETE</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsPositionedDelete()</td>
<td>Is some form of outer join supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SCHEMAS_IN DML</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSchemasInDataManipulation()</td>
<td>Can a schema name be used in a data manipulation statement?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SCHEMAS_IN INDEX_DEFN</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSchemasInIndexDefinitions()</td>
<td>Can a schema name be used in an index definition statement?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SCHEMAS_IN RIVILEGE_DEFN</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSchemasInPrivilegeDefinitions()</td>
<td>Can a schema name be used in a privilege definition statement?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SCHEMAS_IN PROCEDURE_CALLS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSchemasInProcedureCalls()</td>
<td>Can a schema name be used in a procedure call statement?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SCHEMAS_IN TABLE_DEFN</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSchemasInTableDefinitions()</td>
<td>Can a schema name be used in a table definition statement?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SELECT_FOR_UPDATE</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSelectForUpdate()</td>
<td>SELECT for UPDATE supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_STORED_PROCEDURES</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsStoredProcedures()</td>
<td>Are stored procedure calls using the stored procedure escape syntax supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SUBQUERIES_N_COMPARISONS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSubqueriesInComparisons()</td>
<td>Are subqueries in comparison expressions supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SUBQUERIES_N_EXISTS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSubqueriesInExists()</td>
<td>Are subqueries in 'exists' expressions supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SUBQUERIES_N_INS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSubqueriesInIns()</td>
<td>Are subqueries in 'in' statements supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_SUBQUERIES_N_QUANTIFIEDS</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsSubqueriesInQuantifieds()</td>
<td>Are subqueries in quantified expressions supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_TABLE_CORRELATION_NAMES</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsTableCorrelationNames()</td>
<td>Are table correlation names supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_TRANSACTION</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsTransactions()</td>
<td>Are transactions supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_UNION</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsUnion()</td>
<td>Is SQL UNION supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPORT_UNION_ALL</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsUnionAll()</td>
<td>Is SQL UNION ALL supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SUPPPORT_POSITIONED_UPDATE</td>
<td>boolean</td>
<td>true/false</td>
<td>supportsPositionedUpdate()</td>
<td>Is some form of outer join supported?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>SYSTEM_FUNCTIONS</td>
<td>String</td>
<td>getSystemFunctions()</td>
<td>List of system functions</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>TABLE_TYPES</td>
<td>String</td>
<td>getTableTypes()</td>
<td>List of supported table types</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>TIME_DATE_FUNCTIONS</td>
<td>String</td>
<td>getTimeDateFunctions()</td>
<td>List of time and date functions</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>URL</td>
<td>String</td>
<td>getURL()</td>
<td>URL of this database</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>USER</td>
<td>User</td>
<td>getUserName()</td>
<td>User of this database</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>USES_LOCAL_FILE_PER_TABLE</td>
<td>boolean</td>
<td>true/false</td>
<td>usesLocalFilePerTable()</td>
<td>Does the database use a file for each table?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>USES_LOCAL_FILES</td>
<td>boolean</td>
<td>true/false</td>
<td>usesLocalFiles()</td>
<td>Does the database store tables in a local file?</td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td>String</td>
<td>getDatabaseProductVersion()</td>
<td>Version for this Database Product</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>NAME</td>
<td>String</td>
<td>getSchema()</td>
<td>Name of Schema</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>Table AUTO_UPDATED_COLUMNS</td>
<td>Bag of Columns</td>
<td>getColumns()</td>
<td>All Columns that are automatically satis ed if a row is updated.</td>
<td></td>
<td>DatabaseMetaData</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>COLUMNS</td>
<td>Bag of Columns</td>
<td>methods</td>
<td>Columns in Table</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------------------------</td>
<td>----------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>------------------------</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>IN_CATALOG</td>
<td>Catalog</td>
<td>getTables()</td>
<td>Table Catalog</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>IN_SCHEMA</td>
<td>Schema</td>
<td>getTables()</td>
<td>Table Schema</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>NAME</td>
<td>String</td>
<td>getTables()</td>
<td>Name of the Table</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>RDF:COMMENTS</td>
<td>String</td>
<td>getTables()</td>
<td>Explanatory Comment on</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>TABLE_PRIVILEGE</td>
<td>Privilege</td>
<td>getTables()</td>
<td>Privileges for this Table</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>TABLE_TYPE</td>
<td>String</td>
<td>getTables()</td>
<td>Type of the Table</td>
<td>DatabaseMetadata</td>
<td></td>
</tr>
</tbody>
</table>

| User           | NAME                    | String         | getUserName()         | Username of this User    | DatabaseMetadata       |
| User           | RIGHTS                  | UserRights     | allProceduresAreCallable(), allTablesAreSelectable() | Rights for this User    | DatabaseMetadata       |
| UserRights     | ARE_PROCEDURES_CALLABLE | boolean        | allProceduresAreCallable() | Whether or not the current user has the rights to call all procedures | DatabaseMetadata       |
| UserRights     | ARE_TABLES_SELECTABLE   | boolean        | allTablesAreSelectable() | Whether or not the current user has the rights to call SELECT statements on all Tables | DatabaseMetadata       |
APPENDIX B

DATABASE CONNECTION POOL PERFORMANCE METRICS

To test the Database Connection Pool's performance and the fact that using a Connection Pool is indeed superior to creating a connection every time, an experiment was conducted on a Pentium II (266) with 128 MB RAM running WIN-NT Workstation 4.0 (SP5).

The Database the testing was performed on was MS-Access 97 (SR2). A tester program was written that spawned 10 threads. Each thread simultaneously executed a SELECT query (select * from test, the table had 6 columns and 250 rows) for a specified amount of time. This program then obtained the ResultSet, stepped through each row and converted each row into a Hashtable with the column names as the keys. This procedure was repeated both with and without the Connection Pool. Without the Pool a connection was created each time a query had to be executed.

The turnaround time (i.e. the time it took for executing the query, stepping through the ResultSet and converting each row into a Hashtable) was recorded for each query. This time was then plotted for both the cases.

Figure 28: Database Connection Pool Performance at a Glance.
## Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Without the Database Connection Pool</th>
<th>With the Database Connection Pool</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time of Run (secs)</strong></td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td><strong>Number of Select Statements</strong></td>
<td>201</td>
<td>351</td>
</tr>
<tr>
<td><strong>Mean Turnaround Time (millis)</strong></td>
<td>24490.40299</td>
<td>13993.73504</td>
</tr>
<tr>
<td><strong>Maximum Turnaround Time (millis)</strong></td>
<td>57783</td>
<td>19828</td>
</tr>
<tr>
<td><strong>Minimum Turnaround Time (millis)</strong></td>
<td>14301</td>
<td>11136</td>
</tr>
<tr>
<td><strong>Median Turnaround Time (millis)</strong></td>
<td>21481</td>
<td>13800</td>
</tr>
<tr>
<td><strong>Number of Threads</strong></td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>
REFERENCES


