Mapping of Oil and Gas Exploration Business Data Entities for Effective Operational Management

Shastri L NIMMAGADDA[1], Heinz DREHER[2] and Amit RUDRA[2]
Wafra Joint Operations, Kuwait Gulf Oil Company, FDD, Kuwait[1]
Curtin Business School, Curtin University of Technology, Perth, Australia[2]

Abstract- **Spatio-temporal data of petroleum resources businesses are heterogeneous in nature with multiple relationships among various entities and attributes. Object oriented (OO) systems provide alternative solutions for handling the complex *exploration* business data entities, where traditional database systems pose serious limitations. *Exploration*, which is a key business object class in any petroleum business environment, is composed of several sub classes, such as *navigation*, *seismic*, *vertical seismic profiling (VSP)*, *well-log* and *reservoir*. Authors classify these typical spatio-temporal data items as classes and sub class objects in the OO modeling. In the present paper, logical entity-relationship (ER) models have been re-written in multi-dimensional and object-oriented models. Syntax of typical exploration data object classes, attributes, operations and their relationships has been described for implementation purposes. This work demonstrates how object class logical data models are flexible and interoperable for fast changing petroleum business situations. Models presented in this paper, guide exploration data managers for effectively managing their operations. An OLAP model discussed in this paper is a pursuit of cost saving detailed exploration for oil and gas prospect investigation in any basin.**

I. INTRODUCTION

OIL and gas industry comprises of intricate hierarchical organizational structure. This causes friction in data and information flow into the organization. Multifaceted organizational structure demands accurate and precise information flow among different operational units of these commercial oil and gas exploration ventures.

Exploration is one of the key operations of oil industry. Nimmagadda et al. [13, 14, 15] describe significance of exploration entity in oil and gas companies. They demonstrate the need of warehousing and data mining technologies in the petroleum companies. In the past, resources company personnel have used databases to store data consisting of only text and numbers. Today, oil and gas exploration professionals store graphics, drawings, photographs, video, sound, voice mails, spreadsheets and other complex objects in their databases. Relational database management systems store these data objects and types with certain limitations. The concept of object, which is the core of all OO systems, is some unit of data (oil and gas

exploration data) along with actions, affecting its behavior (Hoffer et al. [7], Rubin and Goldenberg [20]). A *reservoir* object, for example, could consist of the data relevant to *wells* object (*reservoir* name, type, quality and *production rate*) together with the actions that can take place on *reservoir* object (multi-reservoirs, predicting qualities, *reservoir* extents and thickness for computing the geological and recoverable reserves from petroleum prospects) of oil and gas exploration business. Beaumont and Foster [2] discuss significance of *seismic*, *well-log*, *reservoir* studies in the exploration and exploitation of oil and gas deposits. Authors view typical data items of exploration industry as several object classes as shown in Fig.1.
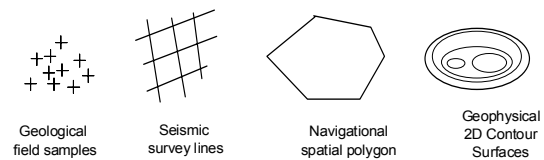

Fig.1. Typical object class data items of exploration industry

For mapping purposes, exploration is considered as a super-type object class (in place of entity). Analogous to entities in ER modeling, sub-type entities are interpreted as sub class objects in object oriented modeling. From super type class objects, sub class objects are derived due to emergence of inherent operational units within major exploration operation. In our study, *navigation*, *seismic*, *VSP*, *well-logging* and *reservoir* are such subclasses of an *exploration* object class. These sub-types are in turn linked with associative classes that are formed due to associativity of common attributes among subclasses. In this paper, all object classes that have common attributes have been linked and integrated to get a metadata model. All the pieces of data required for objects mapping, are uploaded. An integrated multidimensional schema, which is more suited to store exploration data in a warehousing environment, has been proposed. For this purpose, multidimensional view of objects, in the form of OLAP model, has been presented and interpreted. A detailed problem statement follows.

II. PROBLEM STATEMENT

Most databases are relational and or object-oriented - implemented as a set of tables and relationships among their likely associated tables. Although, logical ER models are simple in their visualization and usage, they have certain

1227

limitations in handling complex and heterogeneous exploration data structures and types. The *exploration* data are better understood and managed by object-oriented data modeling, in which sub classes of object classes share common attribute or attributes (of structure) and their behavior. Due to complex nature of oil and gas exploration data, it is proposed to simplify data models by object-oriented data mapping. Initial ER models are presented in logical multidimensional schemas and then rewritten in object oriented data models (Ambler, [1]). *Navigational, seismic*, *VSP*, *well- logging* and *reservoir* data are key participating sub classes of *exploration* class object, in the present data modeling approach. Taking advantage of the unique nature and types of oil and gas exploration data, multidimensional object-oriented data model has been prepared and implemented, integrating individual logical and physical data schemas of *Navig_Seismic_VSP_- WellLog_Reservoir* data structure in a warehouse environment. This data structure representation is innovative and their subclass integration in exploration industry is warranted for successful implementation, in an oil and gas data warehouse environment. This data structure has been analyzed as a part of its implementation in the oil and gas exploration industry and interpreted its future implementation of other classes such as *drilling* and *production* object classes. Shortcomings of present database technologies are briefly presented in the following section.

III. LITERATURE REVIEW

Historical data used in our study, contains period dimension. Khoshaflan and Copeland [10] use identity as a property of an object for modeling purposes. Smith and Zdonik [25] demonstrate a case study, examining the design of an object oriented database schema for Inter-media and comparing with the relational and object-oriented database management system approaches, in which OO models have been made beneficial. Jocobson [9] presents development of object-oriented methodology in an industrial scenario. Kim et al. [11] integrate an object-oriented programming system with a database system. Champeaux et al. [4] depict the process of object-oriented design. Embley [6] provides concepts of objects models. Chung et al. [5] discuss various objects and relational databases, relevant to industry data. Laffra and Stanley [12] investigate the applicability of object-oriented technology in large financial institutions. Price and Demurjian [19] present a framework for reusability measurements, which facilitate large-scale OO design reuse. Box [3] describes object-oriented (OO) technology application in an insurance industry with inherent limitation of flexibility.

Oler and Lindstrom [17] discuss various issues of migrating relational data to ODBMS with a case study of molecular biology. They report on a genetics laboratory database migration experiment, which was facilitated by both organization of the relational data in object-like form and a C++ framework designed to insulate application code from relational artifacts. Zambrono et al. [27] provide an object-oriented perspective to handle continuous fields such as GIS and more specifically in areas like meteorology and environmental applications. Hyunch et al. [8] and Raumbaugh et al. [22] discuss various objects applied to industrial situations. These authors use traditional methods of mapping simple data objects.

Nguyen et al. [16] present an *object-oriented* approach to managing the evolution of system objects at the *logical* level. They demonstrate object-oriented SCM approach by applying it in different development paradigms such as UML-based object oriented software development, architecture-based software development, and Web application development using simple industry scenarios. Seber et al. [24] provide an exhaustive insight on systems design approach for geosciences. Their arguments are limited to the segregation and presentation of different entities involved in geological and geophysical sciences with reference to the middle-eastern basins perspective (Longley et al. [23]). Accessing multidisciplinary and voluminous datasets and how efficiently they can be used, are not well addressed in the previous literature. Methodologies used for investigating present heterogeneous data structure, are discussed in the following section.

IV. MODELING METHODOLOGIES

Data analyst identifies all appropriate relationships between sub-type entities. Attributes that are common to *navigational*, *seismic*, *VSP, well-logging* and *reservoir* sub-type entities, have been used to integrate them, so that the conceptual models derived are true representation of logical data models at later stages. *Exploration,* an operational entity is considered as a super type. The conceptual data modeling linking the entities with their corresponding attributes is used as an initial data model design methodology in the present study. These logical entity-relationship models have been presented for building multidimensional logical models and converting them into object oriented data models.

Data warehouses and data marts are used in a wide range of applications (Hoffer et al. [7]). Business executives in oil and gas industry use the data collected, integrated, preprocessed and stored in data warehouses and data marts, to perform data analysis and make strategic decisions, such as well planning and budgetary proposals for exploration and production. Generating reports and answering predefined queries are key uses of current data warehouse. Performing multidimensional data analysis, such as OLAP and slice and dice operations must address issues of warehouse architecture designs that have been narrated in the Fig.2.
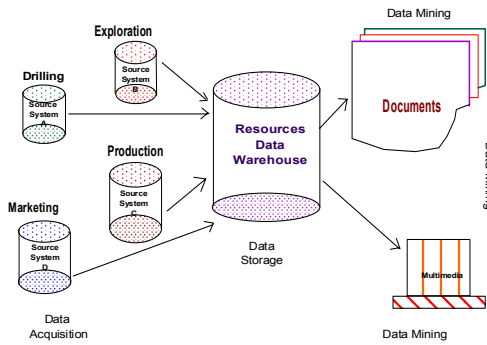
Fig.2. Warehoused petroleum data

To start with, a conceptual model has been constructed using the ER data mapping constructs, such as, entities, attributes and their relationships. As shown Fig. 3, a logical model consisting of *navigational, seismic, VSP* and *well-logging* operations as subclasses of *exploration* class, is presented. Key attributes associated with each and every sub-type entities have been described. Relationships mapped between sub-type entities have been represented with diamond boxes. These relationships evolved due to the business scenario, are represented as associative entities.
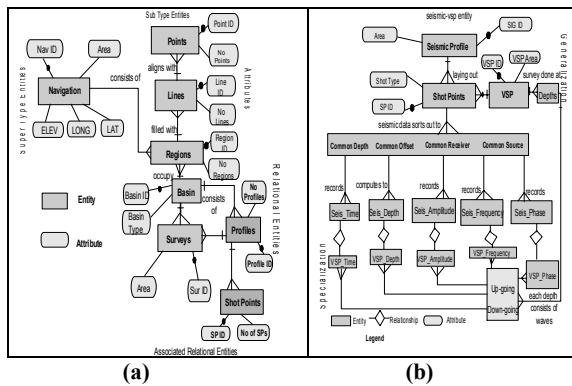

(a)   **(b)**

Fig.3. ER models for (a) navigational-surveys and (b) seismic-VSP entities

Only key entities have been shown in these ER models. Similar ER models constructed for *VSP* (Vertical Seismic Profiling) and *well-logging* entities have been presented in Fig.4. One-to-one, one-to-many and many-to-many relationships have been mapped wherever they exist. How business rules imposed, relevant to exploration data structures, is explained in the following sections.

Enforcing business rules in exploration industry: ER and EER (extended entity-relationship) diagrams are a useful means of setting business rules (Nimmagadda and Rudra [15]). Participation and dis-jointness associated with super-type and sub-type entities or objects are expressions of business rules associated with specialized sub-type relationships.
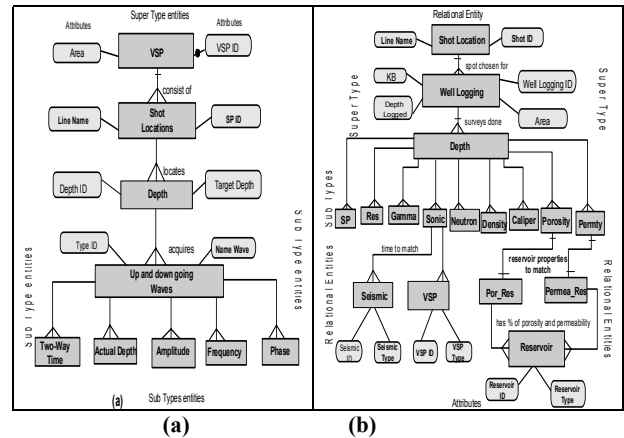

(a)   **(b)**

Fig.4. ER models for (a) VSP and (b) well-logging data entities

Cardinality, optionality, derivation, integrity, action and structural assertions, encapsulation of attributes and operations of oil and gas industry data are key business rules incorporated while mapping complex object data structures. Polymorphism and inheritance are also important concepts of object oriented modeling, applicable to heterogeneous petroleum data structures. In many companies, these business rules have been represented in the form of derivations and algorithms, which enforce organizations and make the object data structures, work as per enforced rules. Logical multidimensional modeling depends on how logical object classes are identified, mapped and tested as per enforced business rules.

Description of relational exploration data structures: As an example, it is demonstrated how these relationships among several entities can be presented in a logical sense using the data structure described in Fig.3 (a). The text description of structure (see ER models shown in Fig. 3) of a relation by a shorthand notation in which the relationships followed by the names of attributes in that relation is represented as:

Navigation (Nav ID, area, longitude, latitude, elevation, Region ID)
Region (Region ID, No_of _Regions, Area_of _Region, Line ID)
Line (Line ID, No_of _Lines, Length, Point ID)
Point (Point ID, No_of _Points, Distance, Shot ID)
Basin (BasinID, No_of_Basins, BasinType, ProfileName, Profile ID)
Profile   (Profile ID, No_of_Profiles, Shot ID)
Shot (Shot ID, Point ID, No_of_Shots, Shot_Distance, ShotLineNumber)

In the above notation, solid lines represent primary key attributes and the dashed lines to foreign key attributes though which the entities have been made interrelated. "Shot" has a composite key consisting of the attribute keys *Shot ID* and *Point ID*. Similar relations can be represented for other data structures shown in Figs 5-8.

After having constructed the logical ER models, DB users check for performance and programming of database structures (Ambler [1]). But *point*, *line* and *area* objects narrated in the *navigational* sub class are difficult to handle

by traditional ER modeling. *Structure region* from *seismic* objects, *oil-water/gas-oil contact region* from *well-logging* object and *net oil-pay region* (reservoir thickness) from *reservoir* objects are difficult to handle by ER traditional modeling. Accessing these *exploration* data, depends on appropriate storing and manipulating these types of data objects. Multidimensional object oriented modeling approach is used to logically map and model these typical data objects. Dimensional data mapping approach (Hoffer et al. [7] and Pratt and Adamski [18]) is one of the logical data model development techniques considered in the present study, since these are used to design exploration data warehouse, accommodating volumes of object oriented multidimensional data structures. All the sub-type classes interpreted from *navigational* and *seismic* operational data object classes have been designated as dimensions as shown in Fig.5, on top of each and every rectangular box (entity).

Fig.6. Dimensional star schema models for (a) VSP and (b) Well-logging objects
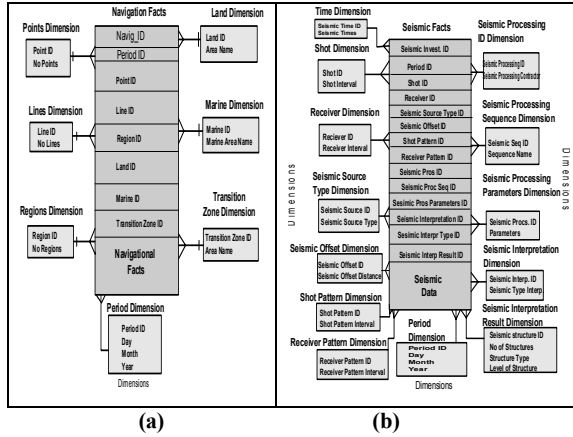

Fig.5. Dimensional star schema models for (a) navigation and (b) seismic data

Primary key attributes of dimensional tables, that are common, are shared and linked with the fact tables through their corresponding foreign key attributes. This procedure is followed for all data structures shown in Figs. 5 and 6.

As shown in Figs. 5 and 6, one can notice that each fact table is surrounded by related dimensional tables, linking them with foreign keys and the corresponding primary keys located in the dimension tables. One-to-many relationships have been mapped in all the tables as shown in Figs. 5 and 6 and prepared implementation models.

Similar logical data models are developed using *surveys* and *navigational* sub class objects. The notation followed, for representing these object class diagrams, is shown in Figs. 7a and b. Entities are represented as dimensions and object classes. Exploration data such as *navigational, seismic, well-logging, VSP* and *reservoir* are represented as objects classes with one-to-many relationships.
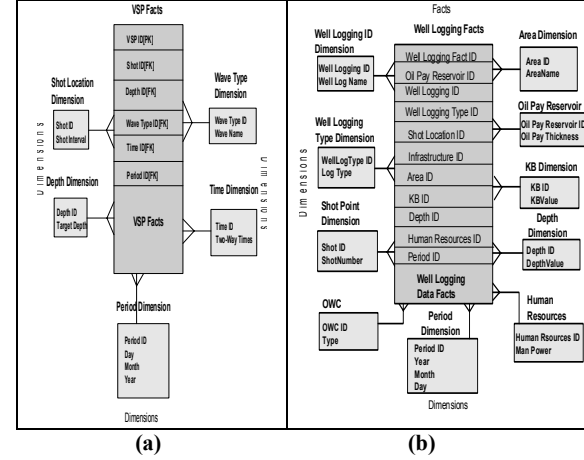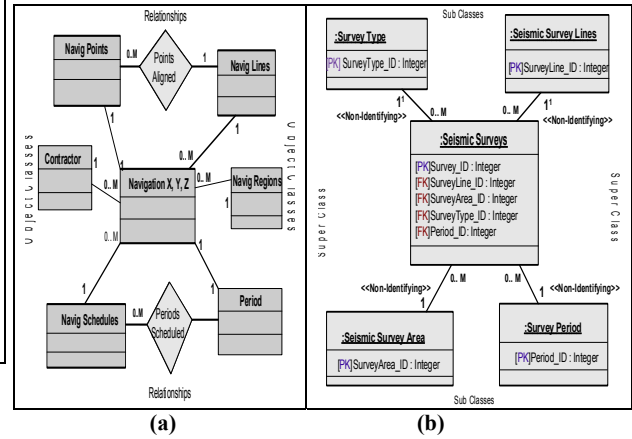

Fig.7. Class diagrams for (a) navigation and (b) seismic objects

V. ANALYSIS AND IMPLEMENTATION

Parallel to the definition of relationship in ER modeling, an association is a 'named relationship among instances of object classes'. The degree of relationship is either unary, binary or ternary or combination of these associations. Petroleum resources industry contains complex data items including spatio-temporal objects (Zambrono et al. [27]). In the object oriented data modeling, object classes and their behavior, share a common set of attributes and behavior. As stated earlier, object oriented databases store complex objects and object schemas constructed using object relational datasets. Hoffer et al. [7] discuss issues relevant to the object relational database design, development and implementation. Several object class models have been generated using sub-type objects, such as *seismic, navigation, wells* as shown in Figs. 5 and 6. Object Definition Language (ODL) is used in building the object oriented database schemas. For oil and gas industry's

2006 IEEE International Conference on Industrial Informatics

exploration data objects, various ODL features are used in building an integrated multidimensional structure model (A OLAP model) as narrated in Fig. 9.

Referential integrity is enforced on the relationships specified in an ODL schema. Another relationship "belongs_to" for the *surveying* class, with the inverse being "conducted" from *survey*. Because, a *survey* that is being conduced, *belongs_to* exactly one *survey*, the destination of the *belongs_to* traversal path is survey, implying that a *surveying ob*ject can be linked to only one *survey* object. In specifying the "one" side of an association relationship, the destination object type is simply specified, and not a collection (e.g., set) of an object type. After having discussed the object data types and one-to-many and many-to-many relationships (1, 1-0..M) established, next stage is to implement them in the exploration business industry. These relationships also affect the super type and associated sub type classes, not only of this object class, but others as well. Relationships have been named among these sub-type object classes as associated classes (similar to relational entities in ER models). As shown in Fig.6, a generalized *exploration* object has been developed. This is a complex object data model, demonstrating how associated classes have been connected to the specialized object classes. Implementation data models are physical representation of data structures in the data warehouses. All identified entities as objects of petroleum geology are integrated in these data models, including the *navigational* data. *Reservoir*, which is an important ingredient of any oil-bearing sedimentary basin, has been related with other objects of super-type *exploration* object. Problems at times, associated with *reservoir*, in particular, the *reservoir* compartmentalization; a significant *infill-drilling* program is needed to maximize recovery, besides the rigorous object mapping analysis. *Structure* object plays a significant role in the entrapment mechanism of associated *reservoirs* as well. Resolving issues of questionable *structure*, such as *fault continuity,* is another possible solution. Distribution of *reservoir* is also controlled by structural factors, where 3D seismic imaging could enhance the *structural pattern* of the oil field. Reservoir quality prediction in sparse *well* grids is another issue, which has recently been addressed by rigorous analysis of *seismic amplitude* objects (Seber et al. [24]) and their characteristics, and proved to be very successful. Superimposition of Reservoir Quality Indicator (RQI) object on the *3D seismic amplitude* object classes (Nimmagadda and Rudra [15] and Longley et al. [23]) (not discussed here) can be used to target the drillable *well* positions. *Pore-fill* expression object can also be mapped and superimposed on *seismic amplitude* objects.

For the purpose of data warehousing design and implementation, object data models are presented in a standard notation as shown in Figs. 7 and 8. *Seismic*, *navigation*, *VSP*, *wells* and *reservoir* and their associated object classes are translated into logical programming modules as discussed in the previous sections.
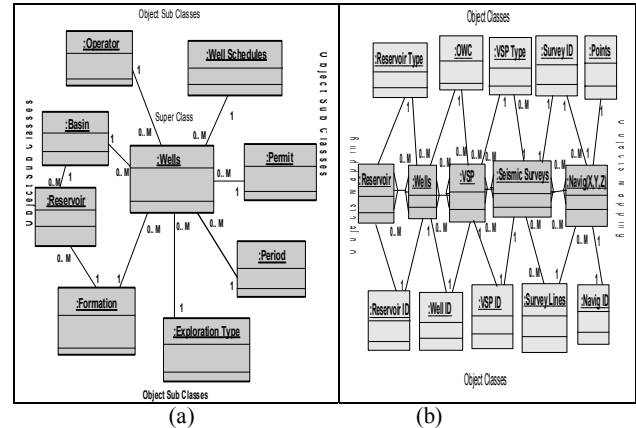

(a)   (b)
Fig.8 Associated class diagrams for (a) wells object and (b) specialized relational objects
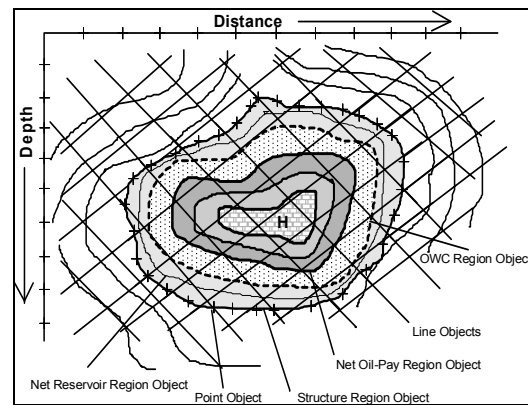

Fig.9 OLAP multidimensional objects data model

Metadata of the current data structure has been used for extracting data views. One of such database view, is processed and generated through mapping algorithms and it is represented as shown in Fig. 9. As demonstrated in Fig 9, multidimensional objects have been extracted from a warehouse and mapped into an integrated data object structure (OLAP model). This is a combined view of several sub-classes and with their associated attributes such as *reservoir* thickness, *oil-water contact* (OWC, an attribute derived from composite data structure, such as *Navig_Seismic_VSP_WellLog_Reservoir*), and *net oil pay* thickness. In other words, all the sub-class objects data such as *points*, *seismic lines* and *regions* have been uploaded and integrated to a single object data model, linking all their respective attributes from multiple object classes. One can interpret this model with an inference that all the ingredients of *oil play*, such as, *structure* and *reservoir* are present in a hydrocarbon province. One can conclude that this is the best drillable exploratory location, which can minimize the risk of *exploration* and also optimize the economics involved in drilling programmes.

VI. FUTURE SCOPE AND CONCLUSIONS

Exploration is one of key business operations of oil and gas business industry. Sub-type entities and objects have been interpreted under super-type exploration business entity. Similar sub-type entities and class objects can be interpreted from other business entities such as *drilling*, *production*, *marketing* and engineering business operations. The present study provides an immense future scope, particularly realizing the complexity and heterogeneity of oil and gas business operations. Object-oriented database programming languages facilitate presentation of oil and gas objects in classes and sub-classes. *Seismic, navigation, VSP, wells* and *reservoir* and their associated object classes are translated into logical programming modules for generating an OLAP model.

Warehouse catering the ODBMS, stores details of persistent exploration data objects and it is the database extension of OO concepts. ODBMS is a synergistic approach for petroleum industry, handling the OO programming languages (OOPL). Clearly, applications that must handle complex heterogeneous structures consisting of data objects are well suited by ODBMS. ODBMS manages data sharing, concurrent data access, and recovery control. Object relational approach is likely to continue to dominate in the market in the years to come. Future research of object relational approach in the oil and gas industry is in the direction of simulating object-oriented layer models and their integration, keeping in view the core concepts of data warehousing and data mining technologies.

Models presented in the present study can further be refined using the ontology base approach. Fine-grained heterogeneous petroleum data structures will facilitate the petroleum knowledge builders (interpreters) in petroleum industry for detailed exploration and development of oil and gas resources of any petroleum bearing sedimentary basin.

VII. REFERENCES

[1] Ambler, S.W. Mapping objects to relational databases: Ronin International, 2000
http://www.AmbySoft.com/mappingObjects.pdf
[2] Beaumont, E.A and Foster, N.H. Exploring for Oil & Gas Traps, *AAPG* Publications of Millennium Edition, 1999.
[3] Box, N. OO Technology for the insurance industry, *OOPSLA*, ACM Press, pp. 110-114, 1997.
[4] Champeaux, D.E. Lea, D. and Faure, P. The process of object oriented design, *OOPSLA*, ACM Press, pp. 45-62, 1992.
[5] Chung, J.Y, Lin, Y.J and Chang, D.T. Objects and relational databases, addendum to the proceedings of *OOPSLA*, pp. 165-169, 1995.
[6] Embley, D.W. Understanding object-model concepts, Addendum to the proceedings of *OOPSLA*, pp. 135-138, 1993.
[7] Hoffer, J.A., Presscot, M.B. and McFadden, F.R. Modern Database Management, 7th Edition, Prentice Hall, 260-520p, 2005.
[8] Huynh, T.N. Mangisengi, O. and Tjoa, A.M. Metadata for Object-Relational Data Warehouse., *Proceedings of the International Workshop on Design and Management of Data Warehouses.*, Stockholm, Sweden., June 5-6, 2000.
[9] Jocobson, I., Object oriented development in an industrial environment, *OOPSLA*, Oct. 4-8, pp. 183-191, 1987.
[10] Khoshaflan, N.S. and Copeland, G. P. Object Identity, *OOPSLA* Proceedings, ACM Press, pp.406-416, 1986.
[11] Kim, W., Ballou, N. Chou, H. Garza, J.F, Woelk, D and Bannerjee, A. Integrating an object oriented programming system with a database system, *OOPSLA*, Oct. 25-30, pp. 142-152, 1988.
[12] Laffra, C. and Stanley, M. OO Technology in large financial institutions, Addendum to the Proceedings of *OOPSLA*, ACM Press, pp. 213-217, 1995.
[13] Nimmagadda, S.L., Dreher, H. and Rudra, A. Data Warehouse Structuring Methodologies for Efficient Mining of Western Australian Petroleum Data Sources, published in *the proceedings of IEEE conference of INDIN 05 held in Perth, Australia,* 2005a.
[14] Nimmagadda, S.L., Dreher, H. and Rudra, A. Warehousing of Object Oriented Petroleum Data for Knowledge Mapping, published in *the proceedings of 5th International Conference of IBIMA*, Cairo, Egypt, 2005b.
[15] Nimmagadda, S.L. and Rudra, A. ER and EER Data Mapping Approaches for Integrating Petroleum Exploration and Production Business Data Entities for Effective Data Mining, published in *proceedings of Kuwait International Petroleum Conference & Exhibition (KIPCE2005)*, Kuwait, 2005c.
[16] Nguyen, T.N. Munson, E.V. and Boyland, J.T. Object-oriented structural software configuration management, *OOPSLA*, pp. 35-36, 2004.
[17] Oler, J. and Lindstrom, G. Migrating relational data to an ODBMS: Strategies and Lessons from a molecular biology experience, *OOPSLA*, ACM press, pp. 243-252, 1997.
[18] Pratt, J.P and Adamski, J.J. Concepts of database management, 3rd Edition, Excellence in Information Systems, Thomson Learning, p.253-275, 2000.
[19] Price, W.M. and Demurjian, S.A. Analyzing and measuring the reusability in object oriented design, *OOPSLA*, ACM Press, pp. 22-33, 1997.
[20] Rubin, K.S and Goldenberg, A. Object behavior analysis, Communications of the ACM, Vol. 35 (9), pp.48-62, 1992.
[21] Rudra, A. and Nimmagadda, S.L. Roles of multidimensionality and granularity in data mining of warehoused Australian resources data, *Proceedings of the 38th Hawaii International Conference on Information System Sciences*, Hawaii, USA, 2005.
[22] Rumbaugh, T. M. Blaha, W, Premarlani, F. Eddy, and Lorensen, W. Object-Oriented Modeling and Design, Englewood Cliffs, NJ: Prentice Hall, 1991.
[23] Longley, I.M. Bradshaw, M.T. and Hebberger, J. *Australian petroleum provinces of the 21st century*, in Downey, M.W. Threet, J.C. Morgan, W.A., 2001, Petroleum provinces of the 21st century, *AAPG Memoir 74*, pp.287-317, 2001.
[24] Seber, D. Steer, D., Sandvol, E. and Sandvol, C. Design and development of Information Systems for the Geosciences: An application to the middle-east, *GeoArabia*, Vol. 5(2), Gulf Petrolink, Bahrain, 2000.
[25] Smith, K.E. and Zdonik, S.B. Intermedia: A Case Study of the Differences Between Relational and Object-Oriented Database Systems, *OOPSLA*, Oct. 4-8, ACM Press, pp. 452-465, 1987.
[26] Watterson, K. "When it comes to choosing a database, the object is value", *Datamation*, pp.100-107, 1998.
[27] Zambrono, A., Polasek, L. Gordillo, S, Bazzocco, J., Tilcara. An OO perspective to handle continuous fields in GIS, *OOPSLA*, ACM Press, pp.185-186, 2000.