Analysis and Identification of Data Heterogeneity on Learning Environment Using Ontology Knowledge

Arda Yunianta
Faculty of Computing,
University Teknologi Malaysia, Malaysia.
Faculty of Information Technology and Communication,
Mulawarman University, Indonesia.
arda.mti08@mail.ugm.ac.id

Norazah Yusof
Faculty of Computing and Information Technology,
King Abdul aziz University, Rabigh, Saudi Arabia.
Faculty of Computing,
University Teknologi Malaysia, Malaysia.
norazah64@gmail.com

Abdul Aziz, Nataniel Dengen
Faculty of Information Technology and Communication,
Mulawarman University, Indonesia.
azizlatte@yahoo.co.id, ndengen@gmail.com

Mohd Shahizan Othman
Faculty of Computing,
University Teknologi Malaysia, Malaysia.
shahizan@fsksm.utm.my

***Abstract*— Heterogeneity on learning environment is about different data and applications to support a learning process in education institutions. Distributed and various systems on learning environment is the current issues to produce big and heterogeneity data problem. A lot of relationships are formed between elements on learning environment. The element on learning environment consists of learning data, learning applications, data sources, learning concept, and data heterogeneity aspect on learning environment. These elements are interrelated and produce complex relationship between each other. A complex relationship problem between elements on learning environment makes a process of analysis and identification difficult to be done. Existing method to drawing this heterogeneity problem make confuse and misunderstanding readers. To solved this problem, researcher using ontology knowledge to describe and draw a semantic relationship that represent the complexity of data relationship on learning environment. The result of this analysis is to develop ontology knowledge to solve complexity relationship on learning environment, and also to help reader's better understanding the complex relationship between elements on learning environment.**

Keywords—learning environment; data heterogeneity; ontology knowledge; semantic approach

I. INTRODUCTION

Data on learning environment is increasingly grown up and becoming more valued and learning process is a part of our daily activities in our life [1]. Implementation of Electronic Learning on learning environments [2] [3] are becoming popular and very important in today's scenario because of their flexibility, convenience and accessibility to support learning activities in traditional learning process.

There is numerous and various application system on learning environments from different function and with specific purpose, this is usually known as heterogeneity on learning environment. The heterogeneity may be the difference in: User interface, Platform, Application system, Database system, Data representation etc.

Data heterogeneity on learning environment is about different data representation and types of information or data in different and numerous applications to support a learning process in education institutions [4]. Different applications are develop for specific purposes based on function and feature that included on that applications [1]. A lot of applications developed on learning environment, such as teaching and learning online application called e-learning, student grading system, student management system, student payment system and etc. In this paper, researchers are using UTM (Universiti Teknologi Malaysia) as a case study to analyze the data heterogeneity problem on university environment. With numerous applications that develop with various system and database schema, produces a big data with heterogeneity problem on that environment.

Different application system with numerous and heterogeneity information, data sources, databases system and data representation makes communication and integration process between this applications difficult to implemented [5] [6]. Various database systems that develop under numerous applications on learning environment support heterogeneity problem on database system aspect [5]. Data representation aspect also support heterogeneity problem on learning environment. Because, different applications are developed on

The authors would like to thank to Fakultas Teknologi Informasi dan Komunikasi, Mulawarman University for their financial support to submit this paper into this conference or journal.

learning environment, they have difference data representation schema on database system.

In the learning process conduct by students and lecturers they have reference sources. The reference sources is not only from lecturer slides but there are a lot of learning material from the other sources such as journal papers, website pages, books, articles, Instagram, YouTube and the other external sources [7]. These learning sources relate with data types and data structures element on data heterogeneity problem.

Learning data on data heterogeneity problem is about learning knowledge that contain on teaching and learning experiences, assessment task conducted by lecturers to students, and learning outcome that have two perspectives are from course outline and students result. These three elements relate with constructive alignment concept that explain a relationships between three elements. Constructive alignment is the one of educational model [8], based upon constructive learning theory (constructivism) and aligned with curriculum. Constructive alignment aims to enhance student outcomes, and this model has been successfully applied to many domains and linked to positive effect to the student results, there is little research related to its application to introductory programming [9].

All the elements that exist in the data heterogeneity on learning environment relate with learning model. With the fundamental goal to provide tutorial service that can support the learning process [10]. Learning model is a part of Intelligent Tutoring Systems concept which uses interface as a way to interact with user learner. Modelling user in e- learning is largely in research; learning model is an essential component in adaptive e-learning systems [1] to support learning process. Therefore, the learning models are used to adapting and modify the interaction between electronic learning system and learner to suit the needs of individual learner to get better learner result [1].

The aim in this paper is to analyze and to identify what kind of data heterogeneity problem that be happen on learning environment, what are the semantic relationship relate between elements on data heterogeneity on learning environment. After analyzing and identification process, the main contribution of this research is to develop ontology knowledge to describe what kind of data heterogeneity problem on learning environment and to get better reader understanding with ontology viewing that contain semantic relationship between elements on learning environment.

This paper is divided into several stages, the first is to analyze and identify what is the elements data heterogeneity on learning environment. The next step is to create semantic relationship between all of elements on learning environment. And the final result is to develop ontology knowledge to draw the complexity that happen on data heterogeneity problem on learning environment with semantic relationship.

II. Literature Review

A. *Constructive Alignment*

Constructive alignment is an educational model to conduct learning process. Two aspect are built on constructive alignment is a constructive aspect refers to the idea that students construct the meaning on learning process through relevant learning activities. The meaning is not something imparted or transmitted from teacher to learner, but is something learners have to create from their creative thinking. Teaching is simply a catalyst for learning. The second aspect is an alignment aspect, alignment aspect is refers to what the teacher does, which is to create a learning environment that supports the learning process appropriate to achieving the desired learning outcomes. The key is that the components in the teaching system, especially the teaching methods used and the assessment tasks are aligned with the learning activities assumed in the intended outcomes. The learner is a peoples who want get better understanding to solve learning problem that given to them [12] [13].

To conduct a better learning process trough constructive alignment is how to perform all assessment tasks and teaching learning process (included content and methods) must be linked to the desired unit of study learning outcomes. Fig1 shows the interrelationship between the three elements on constructive alignment concept [8].

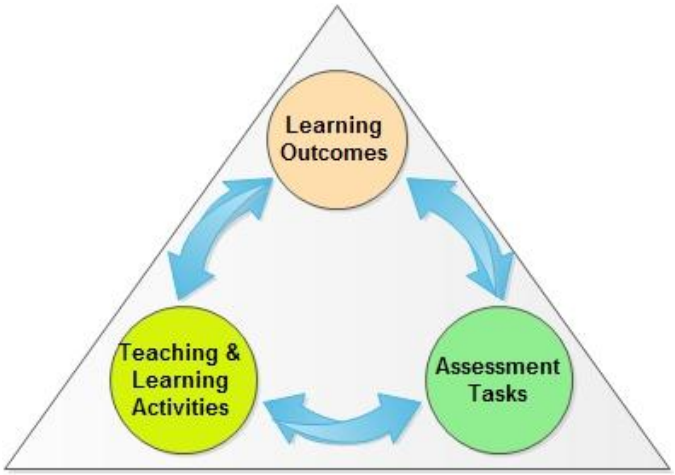

Fig. 1. Constructive Alignment Concept

The important thing on a constructive alignment is to get balance between teaching and learning activities, assessment tasks activities and learning outcomes to be achieved. A Constructive alignment is an educational model [9], this concept based upon constructive learning theory (constructivism) and aligned curriculum. The aim of this concept is get better students' performance to enhance student outcomes. While the model has been implemented on many domains and linked to positive effect to the student results, also for learning environment this concept really closed with all of activities on learning process [10].

Constructive Alignment concept is comprised by a set of principles that can be used to devising Teaching and Learning Activities that help in achieving the Intended Learning Outcomes. This is accomplished by carefully aligning and

learning assessments to support the students to fulfill the Intended Learning Outcomes [8] [12] [13].

B. Learning Model

Learning model is a part of Intelligent Tutoring Systems concept that completed with interface as a bridge to communicate with user learner. This concept was born on 1982 from researcher name are Sleeman and Brown. After this concept growing up and began much used, in the second era on 1987, Wenger's Artificial Intelligence and Tutoring Systems which established what would become the so-called "traditional ITS architecture" with its four components: domain, student, tutor and user interface [10]. Intelligent tutoring systems (ITSs) must be equipped with an explicit representation of the domain knowledge that is the subject of the learning activity to enable learners to acquire knowledge and develop skills in a specific domain. It must also be equipped with the mechanisms by which can be to solve problems in the domain through acquired knowledge and better skill development, because this is a main purpose of intelligent tutoring systems [15].

The purpose of the learning process is how to provide better personalized service and teaching materials for learners to achieved better learners result. Learners cannot learn with the same model, because every learner has a different attitude and it cannot satisfy the learner with his personalized need. But if the learner model makes initial estimations of the new learner's knowledge level and his personalized characteristics before his study, the problem of personalized need will be solved [14]. Fig2 shows the interrelationship three elements on learning model as a main part of intelligent tutoring system [15].

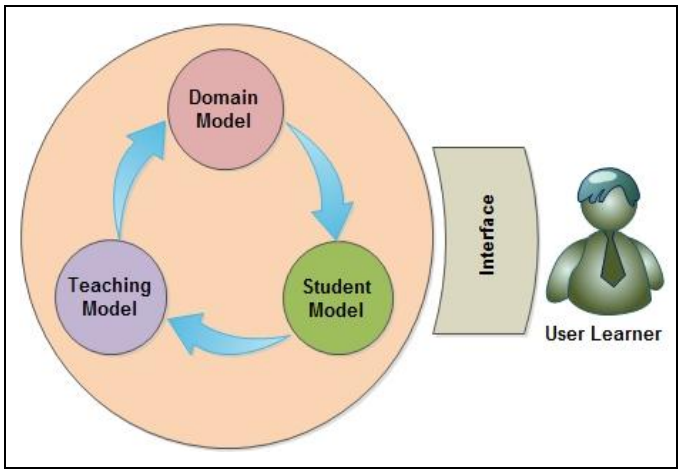

Fig. 2. Learning Model as a Part of Intelligent Tutoring Systems

The domain model is also called expert knowledge. This model contains the problem solving strategies, rules and concept of the domain to be learned. It can fulfill several roles: as a source of expert knowledge, a standard for evaluating the student's performance or for detecting errors, etc. The domain model can organized to be a curriculum, learning structure that including all the learning knowledge elements and linked together according to pedagogical sequences. Each learning knowledge unit on curriculum can be more or less detailed and the curriculum can be practically organized in a dynamic model, according to various structures such as hierarchies, frames, production rules, ontology and semantic networks [10].

The student model is the core component of an ITS. Ideally, it should contain as much knowledge as possible about the students cognitive that support learning process to get better understanding of student's perspective to solve learning problem to get better students result [10].

The tutoring model receives input from the domain and student models and makes decisions about tutoring strategies and actions. Based on principled knowledge, it must make such decisions as whether or not to intervene, and if so, when and how. Content and delivery planning are also part of the tutoring model's functions. Tutoring decisions would ideally be reflected in different forms of interaction with the student: Socratic dialogs, hints, feedback from the system, etc.

More generally, learner and tutor interactions usually occur through the learning interface, also known as the communication or interface component. This component gives access to the domain knowledge elements through multiple forms to interact with user learner.

C. Heterogeneity on Learning Environment

Nowadays, development system on learning environment growing fast. A lot of application system developed to help learning process in some education institutions. Every application is develop for specific purposes with several facilities. Each application is developing with different function, purpose and with different developer produce heterogeneity aspect on the application perspective. Application developer under different programming language, system architecture, model, interface, platform and databases system make communication and integration process more difficult.

Identification for the heterogeneity application on learning environment based on specific function and purposes can be described into several system. Start from teaching and learning online application called e-learning, student grading system, student management system, student payment system until students and lecturer's evaluation system.

Database system aspect also completed the data heterogeneity problem on learning environment. Some applications are develop under the same database system and the other applications are developed with different database system. Researchers are identifying a several database system to develop an application, there are Oracle database system, MySQL database system, SQL Server, and Postgre SQL. Different data are saved in the database system also have different schemas to develop the database system this is produce a different data representation format.

Kinds of data and information that handle by application also have a heterogeneity aspect on that field. There are kind of data structure and a kind of data types. These two aspects have a direct relationship with learning information from external sources. On data structure divided into three, structured data, semi-structured data and unstructured data. And for the data types researchers are divided into five types

of data, there are database data, web/web services data, text data, images data and multimedia data. These all of elements are related with learning information from external sources. Learning information from external sources are from websites, journal papers, books, Instagram and YouTube.

III. Semantic Relationships Between Elements On Learning Environment

After analysis and identification process, the next step is to draw the complexity relationship between elements on learning environment. For the main elements of data heterogeneity aspect, researcher divided into five elements. The five elements are data and information Learning (Learning knowledge), data types, data structure, database system and data representation schema.

These five elements will be related with the other four main elements on learning environment there are constructive alignment concept, learning information from external sources, learning applications, and learning model. These five main elements will be split into several smaller elements.

Researcher will draw this complexity relationship with basic images to shows to the reader what a mess relationship on this problem. After draw and shows the basic image, researcher will identify and describe what the suitable relationship to use in that element. Researcher will draw the relationship using ontology approach to show semantic relationship between elements on learning environment and to get better understanding from readers to understand complexity relationship on data heterogeneity on learning environment. Fig3 shows the complexity relationship between all elements on data heterogeneity on learning environment.

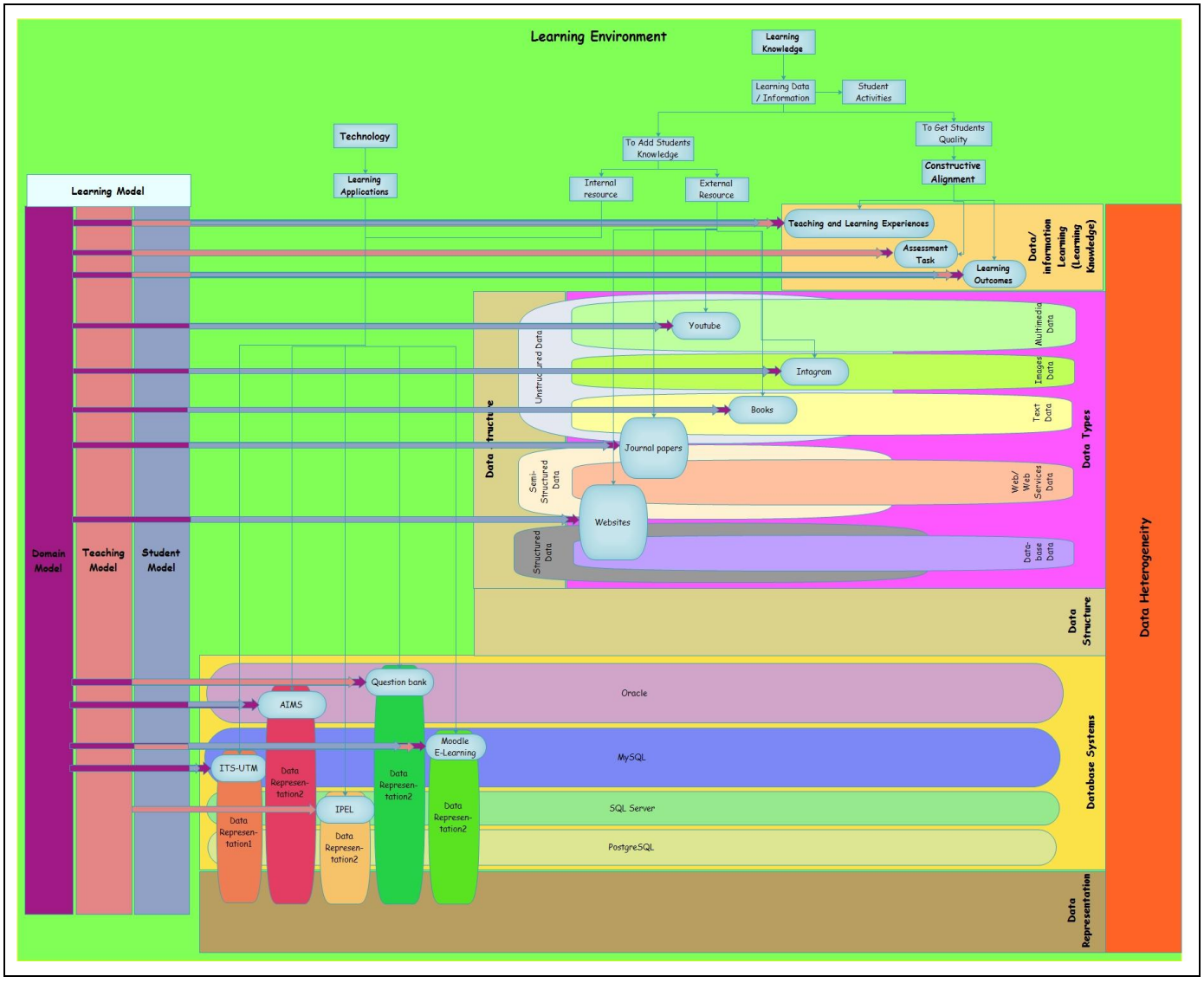

Fig. 3. Data Heterogeneity on Learning Environment

Data heterogeneity complexity shows on Fig3 as a basic representation to shows all of elements on data heterogeneity on learning environment and relationship between elements. After this researcher will use a n ontology approach to get better representation from complexity relationship between elements on learning environment. Semantic relationship created to provide better relationship between elements

because semantic relationship gives a name of relationship between elements.

Fig. 4. Constructive Alignment Semantic Relationship

All of elements on Constructive alignment from learning environment perspective is a same with elements on Learning knowledge from data heterogeneity perspective. Ontology knowledge will shows more detail semantic relationship between these two main elements.

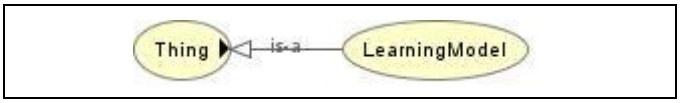

Fig. 5. Learning Model Semantic Relationship

Learning model as a main element will have three child elements are domain model, teaching model and student model.

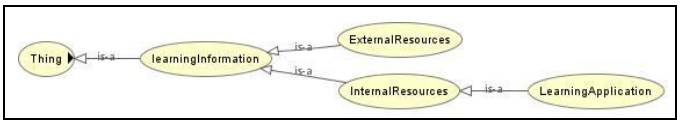

Fig. 6. Learning Information Semantic Relationship

Learning information is a learning sources to add a student knowledge. Learning information have two child elements are external resource and internal resource. Internal resource from learning information has relationship with learning application.

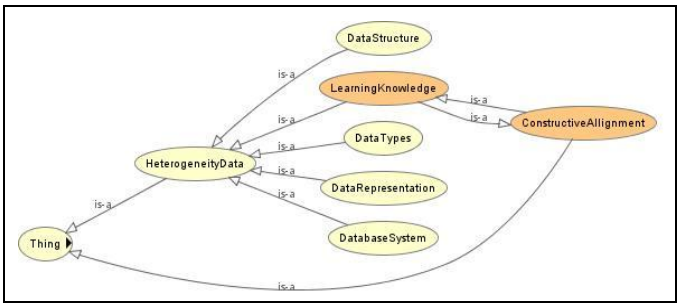

Fig. 7. Data Heterogeneity Semantic Relationship

Data heterogeneity have five main elements are data representation, data structure, data types, learning knowledge and database system. Learning knowledge will have relationship with constructive alignment element.

IV. Ontology Knowledge Data Heterogeneity on Learning Environment

Analysis and identification process to describe a complexity on data heterogeneity on learning environment have shown on previous chapter on this paper. After create a semantic relationship between elements on learning environment, the next step is to developing ontology to shows more derailed all of semantic relationship on data heterogeneity on learning environment.

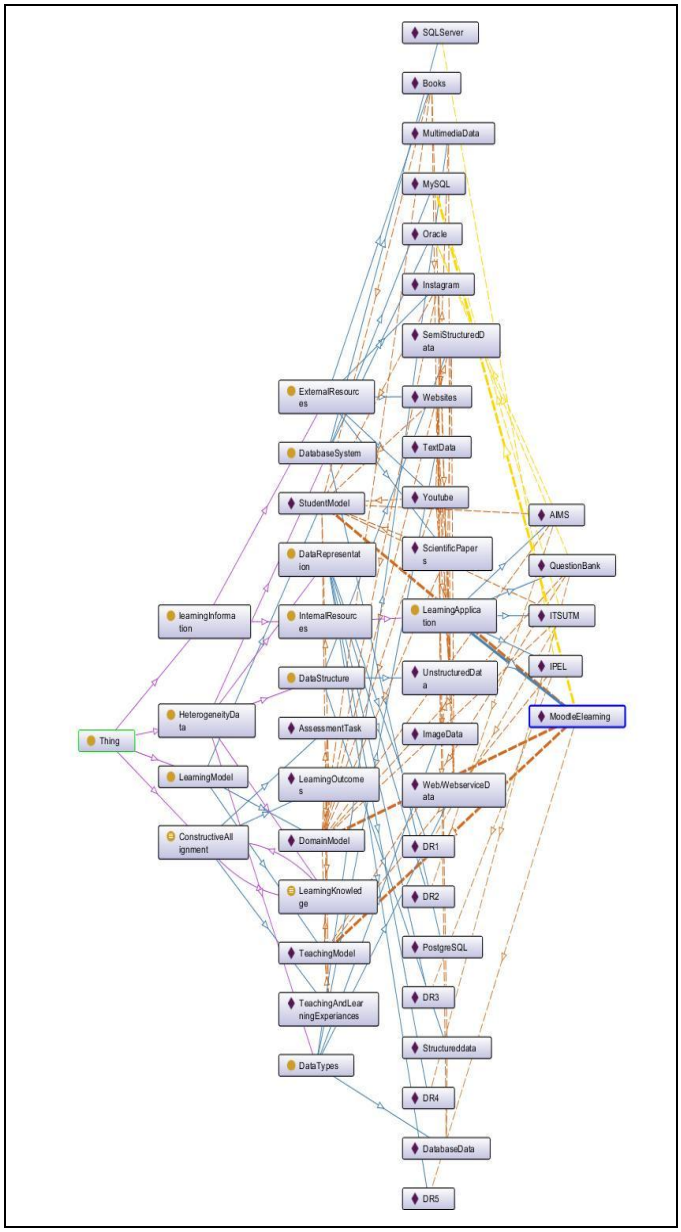

Fig. 8. Ontology Knowledge Data Heterogeneity on Learning Environment

Ontology knowledge that shows on Fig 8 is a complete semantic relationship between all of elements on learning environment. Reader can see more detailed a main elements and child elements from this ontology. Main elements on ontology knowledge is described as a Class and sub Class, while for a Child elements on ontology knowledge is described as a instances or individuals. The class and subclass on ontology will have relationship with the other class or subclass. This example can be seen on constructive alignment and learning knowledge. All of instance that exist in constructive alignment also exist in learning knowledge.

That's mean teaching and learning experiences, assessment tasks, and learning outcomes owned by two different classes.

Every class, subclass and instance on ontology possible to have more than one relationship. On this ontology knowledge researcher describe a four main class under root class (Thing), there are learning information, heterogeneity data, learning model and constructive alignment class. From four class there are two class that have subclass as a detail of each class. Heterogeneity class has five sub classes are database system, data representation, data structure, data types and learning knowledge. With some condition learning knowledge is a constructive alignment. That mean all of instance are exist on learning knowledge also exist on constructive alignment.

Class learning information have a two subclasses are external sources and internal sources. Because of internal resources have semantic relationship with learning application, researcher describe a learning application as a subclass of internal resources.

Complexity of relationship on data heterogeneity on learning environment id describe as an object properties on ontology knowledge. Researchers have described four object properties on ontology knowledge, there are implement, isA, represent and using. Class that using all of object properties is a learning application class. Because on an application class they have five subclasses are AIMS, Question bank, ITS UTM, IPEL and Moodle e-learning. Researchers are using UTM (Universiti Teknologi Malaysia) as a case study to analyze the data heterogeneity problem on university environment. The five applications is an representation from a student grading system, student evaluation system, student management system, student and teaching evaluation system, teaching and learning online application system.

Every instance on learning application has a more than one object property. AIMS have four object property are represent DR2, isA DomainModel, using Oracle and isA StudentModel. That's mean this instance have a relationship with DR2 instance in DataRepresentation class, DomainModel instance in LearningModel class, Oracle instance in DatabaseSystem subclass, and StudentModel instance in LearningModel class. IPEL have three object property, there are using SQLServer, isA TeachingModel and represent DR3. That's mean this instance have relationship eith SQLServer instance in subclass DatabaseSystem, TeachingModel instance in LearningModel class and DR3 inctance in DataRepresentation subclass.

ITSUTM have four object property, there are isA DomainModel, using MySQL, isA StudentModel and represent DR1. That's mean this instance have relationship with DomainModel instance in LearningModel class, MySQL inctance in DatabaseSystem subclass, StudentModel instance in LearningModel class and DR1 instance in DataRepresentation subclass. Questionbank have four object property, there are representDR4, isA DomainModel, suing Oracle and isA TeachingModel. That's mean this instance have relationship with DR4 inctance in DataRepresentation subclass, DomainModel instance in LearningModel class, Oracle instance in DatabaseSystem subclass and TeachingModel inctance in LearningModel class. While for MoodleElarning instance have five object property, there are isA StudentModel, isA teachingModel, isA DomainModel, using MySQL and represent DR5. That's mean this instance have relationship eith TudentModel instance in LearningModel class, TeachingModel instance in learningModel class, DomainModel instance in LearningModel class, MySQL instance in DatabaseSystem subclass and DR5 instance in DataRepresentation subclass.

V. CONCLUSION

A complexity relationship problem on learning environment is very difficult to draw and describe using basic method. There are many things on heterogeneity aspect on learning environment such as elements, aspects and relationship between them. The element on learning environment consists of learning model, learning applications, learning sources, and learning data concept (constructive alignment). While for aspects on learning environment is a data heterogeneity aspect on learning environment. Data heterogeneity aspect on learning environment consists of learning knowledge, data types, data structure, database system and data representation. All of elements and aspects on learning environment are interrelated and produce complex relationship between each other. A complex relationship problem between elements on learning environment makes a process of analysis and identification difficult to be done. Semantic technology through ontology knowledge is a current approach to solve a complexity relationship on some domain. Ontology knowledge using a semantic relationship to handle a numerous and various relationship that exist on that. Ontology knowledge is a better solution to handle data heterogeneity problem on learning environment.

ACKNOWLEDGMENT

The authors would like to thank to Fakultas Teknologi Informasi dan Komunikasi, Mulawarman University for their financial support to submit this paper into this conference or journal.

REFERENCES

[1] Shyamala, R., R. Sunitha, and G. Aghila, Towards Learner Model Sharing Among Heterogeneous E-Learning Environments. International Journal of Engineering Science and Technology (IJEST), 2011. 3(4): p. 2034-2040.

[2] Xiaofei, L., A.E. Saddik, and N.D. Georganas. An implementable architecture of an e-learning system. in Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on. 2003.

[3] Dietinger, T., ASPECTS OF E-LEARNING ENVIRONMENTS, in Institute for Information Processing and Computer Supported New Media (IICM)2003, Graz University of Technology: Austria.

[4] Yunianta, A., et al., Ontology Development to Handle Semantic Relationship between Moodle E-Learning and Question Bank System, in International Conference on Soft Computing and Data Mining2014: Johor, Malaysia.

[5] Huang, O.-r., et al., Application of Ontology-based Automatic ETL in Marine Data Integration, in Symposium on Electrical & Electronics Engineering (EEESYM)2012, IEEE

[6] Wang, C.-C., W.-C. Pai, and N.Y. Yen, A Sharable e-Learning Platform Based on Cloud Computing, in Computer Research and Development (ICCRD), 2011 3rd International Conference on2011. p. 1-5.
[7] Gudanescu, N., Using modern technology for improving learning process at different educational levels. Procedia - Social and Behavioral Sciences, 2010. 2(2): p. 5641-5645.
[8] Biggs, J., Enhancing teaching through constructive alignment. Higher Education, 1996. 32(3): p. 347-364.
[9] Cain, A. and C.J. Woodward. Toward constructive alignment with portfolio assessment for introductory programming. in Teaching, Assessment and Learning for Engineering (TALE), 2012 IEEE International Conference on. 2012.
[10] Nkambou, R., J. Bourdeau, and R. Mizoguchi, Introduction: What Are Intelligent Tutoring Systems, and Why This Book?, in Advances in Intelligent Tutoring Systems, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Editors. 2010, Springer Berlin Heidelberg. p. 1-12.
[11] Vanfretti, L. and M. Farrokhabadi. Implementing constructive alignment theory in a power system analysis course using a consensus model. in e-Learning in Industrial Electronics (ICELIE), 2012 6th IEEE International Conference on. 2012.
[12] J. Biggs, and C. Tang. Teaching for Quality Learning at University. Berkshire, England: Open University Press, 2007.
[13] J. Biggs. Aligning Teaching for Constructing Learning. The Higher education Academy. 2004.
[14] You, D., et al., Flexible Collaborative Learning Model in E-Learning with Personalized Teaching Materials, in Advances in Computer Science, Intelligent System and Environment, D. Jin and S. Lin, Editors. 2011, Springer Berlin Heidelberg. p. 127-131.
[15] v kambou, R., Modeling the Domain: An Introduction to the Expert Module, in Advances in Intelligent Tutoring Systems, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Editors. 2010, Springer Berlin Heidelberg. p. 15-32.Nkambou, R., Modeling the Domain: An Introduction to the Expert Module, in Advances in Intelligent Tutoring Systems, R. Nkambou, J. Bourdeau, and R. Mizoguchi, Editors. 2010, Springer Berlin Heidelberg. p. 15-32.