

การรวมแบบจำลองข้อมูลเป็นวิธีการอันมีประสิทธิภาพวิธีการหนึ่งที่จะทำให้สามารถใช้งานข้อมูลซึ่งถูกจัดเก็บไว้ในแหล่งข้อมูลต่างแหล่งกันรวมกันได้ ด้วยวิธีการนี้ผู้ใช้จะสามารถเข้าถึงข้อมูลทั้งหมดได้ในคราวเดียวกันโดยอาศัยการติดต่อที่จุดเดียว อย่างไรก็ตาม ความต่างแบบของข้อมูลที่น่ามารวมทำให้ผู้รวมจะต้องคำนึงถึงความถูกต้องและความต้องการของข้อมูล ปัจจัยสำคัญที่ส่งผลกระทบต่อการแก้ปัญหาความขัดแย้งของความต่างแบบระหว่างแบบจำลองก็คือความหมายที่ครบถ้วนของข้อมูลที่ได้รับการแสดงไว้ในแบบจำลอง เนื่องจากแบบจำลองข้อมูลเชิงวัตถุแสดงไว้เฉพาะความหมายของข้อมูลในเชิงโครงสร้าง (เช่น คลาส แอททริบิวต์และความสัมพันธ์ระหว่างคลาส) การจะแสดงความหมายอื่นเพิ่มเติม (เช่น ค่าข้อมูล หน่วยข้อมูล รวมถึงคำที่มีความหมายเหมือนกันและคำที่เป็นคำลูกกลุ่มกัน) จำเป็นต้องอาศัยวิธีการอื่นดังเช่น การอธิบายแบบจำลองข้อมูลด้วยออนโทโลยี

วิทยานิพนธ์ฉบับนี้นำเสนอแนวคิดใหม่ในการนำออนโทโลยีมาใช้รวมแบบจำลองข้อมูลเรียกว่า “แนวคิดแบบรวมหลายออนโทโลยี” ซึ่งเป็นวิธีการในการรวมแบบจำลองข้อมูลเชิงวัตถุสองชุดเข้าด้วยกันโดยอาศัยการวิเคราะห์จากออนโทโลยีที่แสดงแทนแบบจำลองข้อมูลแต่ละชุด ภายในงานวิจัยได้กำหนดให้นำออนโทโลยีมาใช้อธิบายความหมายของข้อมูลในแต่ละแบบจำลองข้อมูล จากนั้นออนโทโลยีจะถูกนำไปวิเคราะห์และเปรียบเทียบกันเพื่อตรวจสอบว่ามีความคล้ายคลึงหรือแตกต่างกันมากน้อยเพียงใด แล้วนำผลที่ได้มาสร้างเป็นออนโทโลยีรวม เพื่อจะนำมาใช้ในการสอบถามข้อมูลในระบบสารสนเทศรวม

งานวิจัยนี้ตั้งอยู่บนสมมติฐานที่ว่า หากนำแบบจำลองข้อมูลที่ได้รับการอธิบายและแสดงแทนความหมายในรูปแบบที่เหมาะสมและเพียงพอ มารวมเข้าด้วยกันโดยใช้ขั้นตอนวิธีที่เหมาะสม จะทำให้ได้ระบบสารสนเทศรวมที่ถูกต้องและครบถ้วน

Data model integration is an effective method to interoperate data that reside in different sources for the purpose of providing users with a single point of access to those data. Due to data heterogeneity, data correctness and consistency are significant for integration. Richer semantics of data is a major factor in resolving conflicts among heterogeneous data models. As object-oriented data model represents only schema-based semantics of data (e.g. classes, attributes, and class relationships), an alternative method such as ontology is useful for representing additional semantics (e.g. data values, data units, and synonym and hypernym lists).

This thesis proposes a new approach to an ontology-based integration of data models, called “Integrated Multiple Ontology approach”, which provides a method for integrating two object-oriented data models by using an analysis of their ontologies. In this work, ontology will be used to describe semantics of data in each data model. Then the ontologies are analysed and compared to determine their similarities and differences. The result of the comparison is used to devise an integrated ontology that will enable querying on the integrated information.

This work is based on an assumption that the combination of a good knowledge representation that describes adequate semantics of the data model and a suitable integration algorithm leads to the correct and consistent integrated information system.