

Применение слияния онтологий для интеграции информационных систем в задаче баланса производственных мощностей

Н.Г. Ярушкина¹, А.А. Романов¹, А.А. Филиппов¹, А.Ю. Долгановская¹,
М.С. Григоричева¹

¹Ульяновский государственный технический университет, Северный Венец 32,
Ульяновск, Россия, 432027

Аннотация. В данной статье описывается метод интеграции информационных систем авиастроительного предприятия с системой баланса производственных мощностей на основе слияния онтологий. Для каждого информационного обеспечения, интегрируемых информационных систем, формируется онтологическое представление. Онтологическое представление формируется в процессе анализа структуры реляционной базы данных информационной системы. Интегрирующая модель данных формируется на основе слияния онтологических представлений для каждой информационной системы. Интегрирующая модель данных представляет собой механизм семантической интеграции источников данных.

1. Введение

В рамках работы над автоматизацией процесса баланса производственных мощностей (БПМ) авиастроительного предприятия [1] необходимо учитывать наличие на предприятии множества гетерогенных информационных систем, автоматизирующих различные бизнес-процессы предприятия. Согласованности данных можно достичь путем интеграции системы БПМ с существующим на предприятии информационным обеспечением (ИО). Интеграция данных подразумевает объединение данных, находящихся в различных источниках, и предоставление данных пользователям в унифицированном виде. К основным проблемам интеграции данных относятся:

1. Разнородность (гетерогенность) моделей данных.
2. Автономность – независимость ИС друг от друга.
3. Распределенность – данные могут располагаться в различных сегментах локальной сети предприятия и/или в сети Интернет.
4. Различие формата данных.
5. Различие в представлении значений.
6. Потеря актуальности данных одним из источников.

Таким образом, при организации информационного взаимодействия БПМ с существующим на предприятии ИО возникает необходимость решения следующих методологических задач [2, 3, 4, 5]:

1. Создание интегрирующей модели данных, являющейся основой единого пользовательского интерфейса в системе интеграции.
2. Разработка методов отображения моделей данных и построение отображений в интегрирующую модель для конкретных моделей, поддерживаемых отдельными источниками данных.
3. Интеграция метаданных, используемых в системе источников данных.
4. Преодоление неоднородности источников данных.
5. Разработка механизмов семантической интеграции источников данных.

2. Подход к интеграции информационных систем на основе интегрирующей модели данных

Для решения задачи построения интегрирующей модели данных различных информационных систем (ИС) обычно используются методы Linked Data (принцип связанных данных) [6]. Принцип связанных данных предполагает наличие стандартных средств и механизмов для определения наличия и семантики связей между сущностями, представленными данными. В качестве такого средства используется язык представления знаний OWL [7], позволяющий описывать сущности ПрО и отношения между ними.

Таким образом, язык представления знаний OWL используется в интегрирующей модели данных в качестве единой, объединяющей метамодели данных, опирающейся на общие словари путем предоставления терминов из различных словарей внешних источников.

2.1. Онтологическая модель данных

Для реализации интегрирующей модели данных ИО системы БПМ с ИО существующих на предприятии ИС используются методы онтологического инжиниринга.

Онтология – модель представления знаний некоторой ПрО, состоящая из множества определений основных понятий (имен классов, индивидуалов, свойств и т.д.), а также различного рода семантических связей между ними. В основе онтологии находится словарь терминов, отражающих понятия ПрО, и совокупность правил (аксиом), согласно которым эти термины могут быть скомбинированы для построения достоверных утверждений о состоянии рассматриваемой области в некоторый момент времени [8].

Таким образом, основываясь на онтологическом подходе, онтология модели данных будет иметь вид:

$$O = \langle C, P, L, R \rangle, \quad (1)$$

где $C = \{C_1, C_2, \dots, C_n\}$ – множество классов онтологии модели данных; $P = \{P_1, P_2, \dots, P_m\}$ – множество свойств классов онтологии модели данных; $L = \{L_1, L_2, \dots, L_o\}$ – множество ограничений свойств классов онтологии модели данных; R – множество отношений онтологии модели данных следующего вида:

$$R = \{R_C, R_P, R_L\},$$

где R_C – множество отношений, формирующих иерархию классов онтологии модели данных; R_P – множество отношений, определяющих связь вида «класс-свойство» онтологии модели данных; R_L – множество отношений, определяющих связь вида «свойство-ограничение» онтологии модели данных.

2.2. Отображение модели данных ИО ИС в онтологическое представление модели данных

В настоящее время для организации ИО ИС обычно используются реляционные базы данных (РБД). РБД содержат описание ПрО в виде связанных между собой сущностей (таблиц) [9, 10]. Необходимо разработать метод отображения структуры РБД в онтологическое представление модели данных.

Реляционную модель данных можно представить в виде следующего выражения:

$$RDM = \langle E, H, R \rangle, \quad (2)$$

где $E = \{E_1, E_2, \dots, E_k\}$ – множество сущностей (таблиц) РБД; $E_i = (name, Row, Col)$ – i -я сущность РБД, состоящая из имени и множества строк Row и столбцов Col ; $Col_j = (name, type, constraints)$ – j -й столбец сущности РБД, содержащий свойства: название, тип и набор ограничений; $H = \{H_1, H_2, \dots, H_l\}$ – иерархия сущностей РБД в случае использования функции наследования таблиц, при этом:

$$H_j = E_i D(x) E_h,$$

где E_i, E_h – сущности (таблицы и представления) РБД; $D(x)$ – связь вида «родитель-потомок» между сущностями E_i и E_h РБД; $R = \{R_1, R_2, \dots, R_m\}$ – множество связей РБД, при этом:

$$R_k = E_i \frac{F(x)}{G(x)} E_h,$$

$F(x)$ – связь между сущностями E_i и E_h РБД; $G(x)$ – связь между сущностями E_h и E_i РБД. Функции $F(x)$ и $G(x)$ РБД могут принимать значения: U – единичная связь и N – множественная связь.

Для отображения структуры РБД (выражение 2) в онтологическое представление модели данных (выражение 1) используется следующая функция:

$$F(RDM, O): \{E^{RDM}, H^{RDM}, R^{RDM}\} \rightarrow \{C^O, P^O, L^O, R^O\},$$

где $\{E^{RDM}, H^{RDM}, R^{RDM}\}$ – множество сущностей РБД и отношений между ними (выражение 2); $\{C^O, P^O, L^O, R^O\}$ – множество сущностей онтологии модели данных (выражение 1).

2.3. Формирование интегрирующей модели данных

После отображения структуры РБД каждой из интегрируемых ИС в онтологическое представление модели данных необходимо выполнить процесс формирования интегрирующей модели данных на основе полученных онтологических представлений.

В качестве формального представления интегрирующей модели данных можно использовать определение онтологической системы [11]:

$$\sum^O = \langle O^{META}, O^{IS}, M \rangle,$$

где O^{META} – онтология интегрирующей модели данных (метаонтология); $O^{IS} = \{O_1^{IS}, O_2^{IS}, \dots, O_g^{IS}\}$ – множество онтологических представлений модели данных ИО интегрируемых ИС; M – модель машины вывода.

Для формирования интегрирующей модели данных O^{META} на основе множества онтологических представлений модели данных ИО интегрируемых ИС O^{META} необходимо выполнить следующие шаги:

Шаг 1. Выделение общего понятийного аппарата ПрО

Процесс формирования интегрирующей модели данных O^{META} обеспечивается наличием общей терминологии, и все частные онтологические представлений модели данных ИО интегрируемых ИС должны строиться на основе единого выделенного понятийного аппарата. Понятийный аппарат формируется экспертом ПрО на основе анализа полученных онтологических представлений.

Шаг 2. Создание интегрирующей модели данных O^{META}

На данном шаге в интегрирующую модель данных O^{META} добавляется множество классов верхнего уровня C^{META} , описывающих ИО интегрируемых ИС и используемых в качестве основы в процессе слияния онтологий. Такие классы позволяют составить описание для ИО каждой ИС: название, метод подключения, сведения для авторизации и т.д.

Шаг 3. Формирование иерархии классов интегрирующей модели данных O^{META}

На данном шаге в интегрирующей модели данных O^{META} устанавливается соответствие между иерархиями классов $C^{O_i^{IS}}$ онтологических представлений модели данных ИО интегрируемых ИС из множества $O^{IS} = \{O_1^{IS}, O_2^{IS}, \dots, O_g^{IS}\}$.

Шаг 4. Формирование свойств классов интегрирующей модели данных O^{META}

На данном шаге в интегрирующей модели данных O^{META} устанавливается соответствие между свойствами $P^{O_i^{IS}}$ классов онтологических представлений модели данных ИО интегрируемых ИС из множества $O^{IS} = \{O_1^{IS}, O_2^{IS}, \dots, O_g^{IS}\}$. На данном шаге эксперт принимает решение о том, какие свойства классов должны войти в интегрирующую модель данных O^{META} .

Шаг 5. Определение аксиом классов и свойств и проверка интегрирующей модели данных O^{META} на согласованность

На данном шаге происходит наложение ограничений на свойства и классы интегрирующей модели данных O^{META} с учетом ограничений, присутствующих в онтологических представлениях модели данных ИО интегрируемых ИС из множества $O^{IS} = \{O_1^{IS}, O_2^{IS}, \dots, O_g^{IS}\}$. После этого, полученную интегрирующую модель данных O^{META} необходимо проверить на внутреннюю согласованность с помощью машины вывода M . В данном случае требуется разработка методов проверки выполнения условий ограничений, так как существующие машины вывода не поддерживают работу с такими объектами.

3. Заключение

В данной статье представлена реализация метода интеграции ИС авиастроительного предприятия с системой БПМ. В результате использования принципов связанных данных и онтологического инжиниринга для каждого ИО интегрируемых ИС выполняется процесс отображения структуры РБД ИС в онтологическое представление модели данных. На основе полученных онтологических представлений модели данных ИО интегрируемых ИС, с помощью предложенной методики, формируется интегрирующую модель данных.

4. Литература

- [1] Ярушкина, Н.Г. Интеграция проектных диаграмм и онтологий в задаче балансировки мощностей авиастроительного предприятия / Н.Г. Ярушкина, Т.В. Афанасьева, В.Н. Негода, М.К. Самохвалов, А.М. Наместников, А.А. Романов, Г.Ю. Гуськов // Автоматизация процессов управления. – 2017. – № 4. – С. 85-93.
- [2] Коголовский, М.Р. Методы интеграции данных в информационных системах // Институт проблем рынка РАН [Электронный ресурс]. – Режим доступа: <http://www.ipr-ras.ru/articles/kogalov10-05.pdf> (16.10.2018).
- [3] Кусов, А.А. Проблемы интеграции корпоративных информационных систем [Электронный ресурс]. – Режим доступа: <https://cyberleninka.ru/article/n/problemy-integratsii-korporativnyh-informatsionnyh-sistem> (16.10.2018).
- [4] Морозова, О.А. Интеграция корпоративных информационных систем. – Москва: Финансовый университет, 2014. – С. 8-23.
- [5] Степанов, Д.Ю. Способы интеграции данных корпоративных информационных систем. – Москва: Москва, 2014. – С. 207-213.
- [6] Bizer, C. The Story So Far / C. Bizer, T. Heath, T. Berners-Lee // Home – Tom Heath [Electronic resource]. – Access mode: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf> (15.10.2018).
- [7] OWL 2 Web Ontology Language Document Overview (Second Edition) // World Wide Web Consortium [Electronic resource]. – Access mode: <https://www.w3.org/TR/owl2-overview/> (15.10.2018).
- [8] Gruber, T. Ontology // Tom Gruber [Electronic resource]. – Access mode: <http://tomgruber.org/writing/ontology-in-encyclopedia-of-dbs.pdf> (10.10.2018).

- [9] Poggi, A. Linking data to ontologies / A. Poggi, D. Lembo, D. Calvanese, G. De Giacomo, M. Lenzerini, R. Rosati // Data Semantics. – 2008. – Vol. 4900. – P. 133-173.
- [10] Calvanese, D. Ontop: Answering SPARQL Queries over Relational Databases / D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, G. Xiao // Semantic Web journal [Electronic resource]. – Access mode: <http://www.semantic-web-journal.net/system/files/swj1278.pdf> (11.10.2018).
- [11] Гаврилова, Т.А. Базы знаний интеллектуальных систем / Т.А. Гаврилова, В.Ф. Хорошевский. – Санкт-Петербург: Питер, 2000. – С. 59-98.

Благодарности

1. Исследование выполнено при финансовой поддержке РФФИ и Правительства Ульяновской области в рамках научных проектов № 16-47-732054, № 18-47-732016, № 18-47-730022.
2. Исследование проведено в рамках государственного задания Минобрнауки РФ № 2.1182.2017/4.6 «Разработка методов и средств автоматизации производственно-технологической подготовки агрегатно-сборочного самолетостроительного производства в условиях мульти-продуктовой производственной программы».

Application of the ontology merging in the production capacity planning for the integration of information systems

N.G. Yarushkina¹, A.A. Romanov¹, A.A. Filippov¹, A.U. Dolganovskaya¹,
M.S. Grigoricheva¹

¹Ulyanovsk State Technical University, Severniy Venec 32, Ulyanovsk, Russia, 432027

Abstract. This article describes the method for integration of the aircraft factory information systems with the production capacity planning system based on ontology merging. The process of mapping the database structure into the ontological representation is performed for each information system. An integrated data model is formed based on the ontological representations of each information system database structure. The integrated model is a mechanism for semantic integration of data sources.