

УДК 004.896

## ИНТЕГРАЦИЯ ИНФОРМАЦИОННЫХ СИСТЕМ НА ОСНОВЕ МЕТОДА СЛИЯНИЯ ОНТОЛОГИЙ ПРИ РАСЧЕТЕ БАЛАНСА ПРОИЗВОДСТВЕННЫХ МОЩНОСТЕЙ

© 2018 Н.Г. Ярушкина, А.А. Романов, А.А. Филиппов, А.Ю. Долгановская, М.С. Григоричева

Ульяновский государственный технический университет

Статья поступила в редакцию 01.11.2018

В этой статье представлен подход к интеграции данных на основе интегрирующей модели. Строится онтологическая модель данных, отображение данных из информационных систем в онтологическую модель. Рассматривается поясняющий иллюстративный пример.

*Ключевые слова:* онтология, интеграция информационных систем, модель данных.

*Исследование выполнено при финансовой поддержке РФФИ и Правительства Ульяновской области в рамках научных проектов № 16-47-732054, № 18-47-732016, № 18-47-730022.*

*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»*

### ВВЕДЕНИЕ

При разработке системы, позволяющей планировать балансировку мощностей производства авиастроительного предприятия [1], особое внимание следует обратить на то, что на таком предприятии уже существует фрагментарная автоматизация бизнес-процессов. Существующее информационное обеспечение (ИО) авиастроительного предприятия, как правило, покрывает потребности при выполнении основных бизнес-процессов производства. Однако в настоящее время проявляются следующие проблемы, связанные с динамикой развития ИО предприятия:

1. Увеличение объемов хранимых данных.
2. Расширение границ предметной области (ПрО).
3. Рост численности и количества команд разработчиков и пользователей информационных систем (ИС).

Решение данных проблем требует наложения ограничений по согласованности данных и знаний, находящихся как в хранилищах ИС, так

*Ярушкина Надежда Глебовна, доктор технических наук, профессор, заведующий кафедрой «Информационные системы». E-mail: jng@ulstu.ru*

*Романов Антон Алексеевич, кандидат технических наук, доцент кафедры «Информационные системы». E-mail: romanov73@gmail.com*

*Филиппов Алексей Александрович, кандидат технических наук, доцент кафедры «Информационные системы». E-mail: al.al.filippov@gmail.com*

*Долгановская Александра Юрьевна, аспирант. E-mail: adolganovskaya@mail.ru*

*Григоричева Мария Сергеевна, аспирант. E-mail: gms4295@mail.ru*

и в неструктурированном или неформализованном виде в ряде других источников.

Согласованности можно достичь путем решения следующих методологических проблем интеграции системы оптимизации ресурсов авиастроительного предприятия с существующим на предприятии ИО [2, 3, 4, 5]:

- отсутствие корректного формата или семантического слоя для слияния двух и более несопоставимых наборов данных. Устранение семантических различий между приложениями – одна из наиболее сложных и неформализуемых задач интеграции. Действительно, одна и та же сущность может иметь несколько различных семантик, ограничений и допущений в каждой системе;
- необходимость наличия методологии для документирования таких технических аспектов интеграции, как определение записей, структур данных, интерфейсов и потоков данных в масштабах всей организации;
- необходимость определения правил и методов согласования данных во всех интеграционных проектах.

Интеграция данных подразумевает объединение данных, находящихся в различных источниках, и предоставление данных пользователям в унифицированном виде. Системы интеграции данных могут обеспечивать интеграцию данных на физическом, логическом и семантическом уровнях. Интеграция данных на физическом уровне сводится к преобразованию данных, полученных из различных источников, в определенный унифицированный формат представления. Интеграция данных на логическом уровне подразумевает возможность получения данных,

расположенных в различных источниках, через единую точку входа – метасистему. Интеграция данных на семантическом уровне сводится к наличию единого представления данных с учетом их семантики.

Сложность процесса интеграции различается для каждого из выше представленных уровней интеграции [2, 3, 4, 5]:

- на физическом уровне могут использоваться различные форматы представления данных;
- на логическом уровне может иметь место неоднородность используемых моделей и/или схем данных для различных источников;
- на семантическом уровне различным источникам данных могут соответствовать различные ПрО или фрагменты ПрО, имеющие различный понятийный аппарат.

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 системы оптимизации ресурсов авиастроительного предприятия 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.

## 1. ПОСТАНОВКА ПРОБЛЕМЫ

Цели, которые требуется достичь при реализации механизмов информационного взаимодействия системы оптимизации ресурсов авиастроительного предприятия с существующим на предприятии ИО:

1. Обеспечить полноту и непротиворечивость данных, необходимых для оптимизации

ресурсов производства авиастроительного предприятия.

2. Минимизировать дополнительный ввод информации, требуется максимально эффективно использовать накопленные в хранилищах ИС авиастроительного предприятия сведения.

3. Обеспечить возможность расширения системы. ИО системы оптимизации ресурсов авиастроительного предприятия должно требовать минимальных затрат на процесс сопровождения при добавлении в систему новой функциональности.

Для достижения первой и второй цели следует решить задачу интеграции с существующим ИО авиастроительного предприятия. Основным требованием в данном случае является совмещения интерфейсов и типов данных.

Интеграция позволяет избежать дублированного ввода данных, что в свою очередь сокращает нагрузку на лиц, ответственных за их актуализацию. При этом также устраняется проблема противоречивости данных, поскольку их потоки не будут двунаправленными, т.е. исключается момент синхронизации и одновременной актуализации данных в нескольких ИС. Вместо этого поток данных направляется от производящей стороны – существующие ИС авиастроительного предприятия, в сторону потребителя – системы оптимизации ресурсов авиастроительного предприятия.

Поставленная ранее цель сделать систему расширяемой предусматривает необходимость хранения наиболее полных сведений об объекте ПрО, т.е. информация хранится в виде списков с перечнем атрибутов. При проектировании ИО системы оптимизации ресурсов авиастроительного предприятия следует найти баланс в использовании методов типизации.

## 2. ПОДХОД К ИНТЕГРАЦИИ ИС НА ОСНОВЕ ИНТЕГРИРУЮЩЕЙ МОДЕЛИ ДАННЫХ

Для решения методологической задачи построения интегрирующей модели данных ИС обычно используются методы Linked Data.

Термин Linked Data был введен Тимом Бернерс-Ли и стал известен как «принципы связанных данных» [6]:

1. Использование унифицированных идентификаторов ресурса.
2. Организация доступа к списку ресурсов с применением протокола HTTP.
3. Использование стандартных технологий Semantic Web: RDF, OWL, SWRL, SPARQL.
4. Использование гиперссылок для идентификации не только веб-документов, но и любой сущности ПрО.

Принцип связанных данных предполагает

наличие стандартных средств и механизмов для определения наличия и семантики связей между сущностями, представленными данными. В рамках данного исследования в качестве такого средства используется язык представления знаний OWL [7], позволяющий описывать сущности ПрО и отношения между ними. Язык представления знаний OWL обладает следующими преимуществами [6]:

1. Позволяет связывать сущности ПрО и документы.
2. Ссылки и отношения между сущностями ПрО типизированы.
3. С помощью уникального идентификатора ресурсов существует возможность связывать любыми видами отношений любые сущности ПрО.
4. Каждая сущность ПрО является частью глобальных метаданных, каждая сущность ПрО может быть использована в качестве отправной точки для просмотра всего пространства данных.
5. Возможность установки связи между данными из различных источников.
6. Информация из различных источников может быть объединена путем слияния набора сущностей в один семантический граф.
7. Возможность смешивать термины различных словарей для представления данных.
8. Гибкость структуры модели данных.

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

### 2.1. ОНТОЛОГИЧЕСКАЯ МОДЕЛЬ ДАННЫХ

В рамках данного исследования для реализации интегрирующей модели данных ИО системы оптимизации ресурсов авиастроительного предприятия с ИО существующих на предприятии ИС используются методы онтологического инжиниринга.

Онтология – модель представления знаний определенной 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 ограниченных 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:

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

где  $E_i, E_h$  – сущности (таблицы и представления) РБД;

$F(x)$  – 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).

В процессе отображения структуры РБД в онтологическое представление модели данных необходимо:

*Шаг 1. Формирование классов онтологического представления модели данных*

На основе множества сущностей РБД формируется множество классов  $C$  онтологии  $O$  модели данных  $E_i \rightarrow C_i$ . Количество классов  $C$  должно быть равно количеству сущностей  $E$  РБД. Названием  $i$ -о класса онтологии  $O$  будет являться название  $i$ -й сущности РБД.

*Шаг 2. Формирование свойств классов онтологического представления модели данных*

На основе множества колонок  $Col$   $i$ -й сущности  $E_i$  РБД формируется множество свойств  $P$   $i$ -о класса  $C_i$  онтологии  $O$  модели данных  $Col_j \rightarrow P_j$ . Количество свойств  $i$ -о класса  $C_i$  онтологии  $O$  должно быть равно количеству колонок  $i$ -й сущности  $E_i$  РБД. Названием  $j$ -о свойства  $P_j$  класса будет являться название  $j$ -й колонки  $Col_j$  сущности РБД.

*Шаг 3. Формирование ограничений свойств классов онтологического представления модели данных*

На основе множества колонок  $Col$   $i$ -й сущности  $E_i$  РБД формируется множество ограничений свойств  $L$   $i$ -о класса  $C_i$  онтологии  $O$  модели данных  $Col_k \rightarrow \hat{L}_k$ . Размер полученного множества ограничений  $\hat{L}_k$   $i$ -о класса  $C_i$  онтологии  $O$  должен быть равен количеству ограничений колонок  $i$ -й сущности  $E_i$  РБД. В данном случае возникает ограничение данного подхода, связанное с трудностями отображения ограничений в случае их представления в качестве триггеров или хранимых процедур.

*Шаг 4. Формирование иерархии классов онтологического представления модели данных*

Если в РДБ применяется наследование та-

блиц необходимо сформировать множество отношений  $R_C$  онтологии  $O$  между всеми дочерними и родительскими классами, соответствующим иерархии сущностей РБД  $H \rightarrow R_C$ . В качестве домена  $j$ -о отношения  $R_{jC}$  онтологии  $O$  модели данных указывается ссылка на класс-родитель  $C_{parent}$  онтологии  $O$ . В качестве диапазона отношения  $R_{jC}$  указывается ссылка на класс-потомок  $C_{child}$  онтологии  $O$ .

*Шаг 5. Формирование отношений между классами и свойствами классов онтологического представления модели данных*

На основе множества колонок  $Col$   $i$ -й сущности  $E_i$  и множества связей  $R$  РБД формируется множество отношений  $R_P$  онтологии  $O$ . При этом для каждого  $j$ -о свойства  $P_j$  онтологии  $O$  формируется два вида отношений:

Отношение вида «класс владлец - свойство». В качестве домена  $\bar{k}$ -о отношения  $R_{P\bar{k}}$  онтологии  $O$  модели данных указывается ссылка на  $i$ -й класс  $C_i$  онтологии  $O$ , которому принадлежит данное свойство, а в качестве диапазона ссылка на  $j$ -е свойство  $P_j$  онтологии  $O$ .

Отношение вида «свойство – класс типа данных». В качестве домена  $\bar{k}$ -о отношения  $R_{P\bar{k}}$  онтологии  $O$  модели данных указывается ссылка на  $j$ -е свойства  $P_j$  онтологии  $O$ , а в качестве диапазона ссылка на  $l$ -й класс  $C_l$  онтологии  $O$ , соответствующий  $l$ -й сущности  $E_l$  множества связей  $R$  РБД  $R \rightarrow R_P$ , либо указывается ссылка на  $m$ -й класс  $C_m$  онтологии  $O$ , соответствующий  $m$ -у типу данных РБД  $j$ -й колонки  $Col \rightarrow R_P$ .

*Шаг 6. Формирование отношений между свойствами классами и ограничениями свойств классов онтологического представления модели данных*

На основе множества колонок  $Col$   $i$ -й сущности  $E_i$  РБД формируется множество отношений  $R_L$  онтологии  $O$ . В качестве домена  $j$ -о отношения  $R_{jL}$  онтологии  $O$  модели данных указывается ссылка на  $k$ -е свойство  $P_k$  онтологии  $O$ . В качестве диапазона отношения  $R_{jL}$  указывается ссылка на  $k$ -е ограничения  $L_k$  онтологии  $O$   $Col \rightarrow R_L$ .

Таблица 1 содержит описание сопоставления структурных компонентов РБД с сущностями онтологии  $O$  модели данных.

### 2.3. ФОРМИРОВАНИЕ ИНТЕГРИРУЮЩЕЙ МОДЕЛИ ДАННЫХ

После отображения структуры РБД каждой из интегрируемых ИС в онтологическое пред-

**Таблица 1.** Соответствие компонентов РБД и онтологии модели данных

| Компонент РБД            | Сущность онтологии           |
|--------------------------|------------------------------|
| Таблица                  | Класс                        |
| Представление            | Класс                        |
| Встроенные типы данных   | Класс                        |
| Иерархия таблиц          | Отношение                    |
| Колонка с внешним ключом | Свойство класса<br>Отношение |
| Колонка с данными        | Свойство класса<br>Отношение |
| Ограничение              | Ограничение<br>Отношение     |

ставление 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  $O^{IS}$  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 инте-*

*грирующей модели данных  $O^{META}$*

На данном шаге в интегрирующей модели данных  $O^{META}$  устанавливается соответствие между иерархиями классов  $C^{O_i^S}$  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  $P^{O_i^S}$  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. ИЛЛЮСТРАТИВНЫЙ ПРИМЕР ПРОЦЕССА ФОРМИРОВАНИЯ ИНТЕГРИРУЮЩЕЙ МОДЕЛИ ДАННЫХ

Для иллюстрации процесса формирования интегрирующей модели данных рассмотрим следующий пример: необходимо осуществить интеграцию ИО системы оптимизации ресурсов

**Таблица 2.** Таблица РБД системы источника, реализующая сущность «Оборудование и инструменты»

| Название | Тип данных            | Описание   |
|----------|-----------------------|--|
| t2_ob    | CHAR(200)             | Название   |
| t2_ng    | NUMBER(5)             | Номер группы   |
| t2_nn    | NUMBER(6)             | Порядковый номер   |
| t2_r1    | CHAR                  | Признак 1:<br>0 – оборудование<br>1 – инструмент<br>2 – материал<br>6 – специнструмент |
| t2_r2    | CHAR                  | Признак 2:<br>0 – гостированный<br>1 – специальный                                     |
| t2_r3    | CHAR                  | Признак 3:<br>20 – нет<br>21 – эскиз<br>30 – модель<br>31 – эскиз и модель             |
| t2_p1    | CHAR(2)<br>nullable   | Название параметра 1   |
| t2_z1    | CHAR(8)<br>nullable   | Значение параметра 1   |
| t2_p2    | CHAR(2)<br>nullable   | Название параметра 2   |
| t2_z2    | CHAR(8)<br>nullable   | Значение параметра 2   |
| t2_p3    | CHAR(2)<br>nullable   | Название параметра 3   |
| t2_z3    | CHAR(8)<br>nullable   | Значение параметра 3   |
| t2_gm    | BLOB                  | Геометрическое представление инструмента   |
| up_dt    | DATE                  | Дата последней корректировки   |
| up_us    | CHAR(32)              | Пользователь   |
| t2_dc    | BLOB                  | Прикреплённый документ   |
| t2_vid   | CHAR(4)               | Вид технологической оснастки   |
| t2_doc   | CHAR(100)             | Обозначение документа  |
| t2_prim  | CHAR(100)<br>nullable | Примечания   |
| t2_yyyy  | CHAR(4)               | Год производства   |

авиастроительного предприятия (приемник) с ИС, содержащей данные об оснащённости цехов авиастроительного предприятия (источник).

Система источник содержит в составе своего ИО РБД, в рамках данной БД реализована сущность «Оборудование и инструменты», представленная таблицей 2.

Система приемник содержит в составе своего ИО РБД, в рамках данной БД сущность «Оборудование и инструменты» реализована в виде следующих таблиц:

1. Оборудование (таблица 3).
2. Тип оборудования (таблица 4).
3. Параметры оборудования (таблица 5).
4. Таблица «Оборудование Параметры оборудования», реализующая связь «многие ко многим» для таблиц «Оборудование» и «Параметры оборудования» (таблица 6).

Онтологическое представление  $O^{source}$  сущности «Оборудование и инструменты» (таблица 2) системы источника будет иметь следующий вид:

**Таблица 3.** Таблица «Оборудование» РБД системы приемника

| Название         | Тип данных               | Описание                             |
|------------------|--------------------------|--------------------------------------|
| id               | INTEGER                  | Первичный ключ                       |
| name             | VARCHAR(255)             | Название                             |
| serial_number    | VARCHAR(255)<br>nullable | Серийный номер                       |
| inventory_number | VARCHAR(255)             | Инвентарный номер                    |
| production_date  | DATE                     | Дата производства                    |
| tool_types_id    | INTEGER                  | Ссылка на таблицу «Тип оборудования» |

**Таблица 4.** Таблица «Тип оборудования» РБД системы приемника

| Название | Тип данных   | Описание       |
|----------|--------------|----------------|
| id       | INTEGER      | Первичный ключ |
| name     | VARCHAR(255) | Название       |

**Таблица 5.** Таблица «Параметры оборудования» РБД системы приемника

| Название | Тип данных   | Описание       |
|----------|--------------|----------------|
| id       | INTEGER      | Первичный ключ |
| name     | VARCHAR(255) | Название       |

**Таблица 6.** Таблица «Оборудование Параметры оборудования» РБД системы приемника

| Название          | Тип данных   | Описание                                   |
|-------------------|--------------|--|
| tool_id           | INTEGER      | Ссылка на таблицу «Оборудование»           |
| tool_parameter_id | INTEGER      | Ссылка на таблицу «Параметры оборудования» |
| value             | VARCHAR(255) | Значение параметра                         |

$O^{source} = \langle$

$C = \{ \text{Оборудование и инструменты (Об), CHAR, NUMBER, BLOB, DATE},$

$P = \{ t2\_ob, t2\_ng, t2\_nn, t2\_r1, t2\_r2, t2\_r3, t2\_p1, t2\_z1, t2\_p2, t2\_z2, t2\_p3, t2\_z3, t2\_gm, t2\_p3, t2\_z3, t2\_gm, up\_dt, up\_us, t2\_dc, t2\_vid, t2\_doc, t2\_prim, t2\_yyyy \}$

$L = \{ nullable, \langle \text{length}, 2 \rangle, \langle \text{length}, 4 \rangle, \langle \text{length}, 8 \rangle, \langle \text{length}, 32 \rangle, \langle \text{length}, 100 \rangle, \langle \text{length}, 200 \rangle, \langle \text{length}, 255 \rangle, \langle \text{precision}, 5 \rangle, \langle \text{precision}, 6 \rangle \}$

$R_P = \{ \langle \text{Об}, t2\_ob, \text{CHAR} \rangle, \langle \text{Об}, t2\_ng, \text{NUMBER} \rangle, \langle \text{Об}, t2\_nn, \text{NUMBER} \rangle,$

$\langle \text{Об}, t2\_r1, \text{CHAR} \rangle, \langle \text{Об}, t2\_r2, \text{CHAR} \rangle, \langle \text{Об}, t2\_r3, \text{CHAR} \rangle, \langle \text{Об}, t2\_p1, \text{CHAR} \rangle,$

$\langle \text{Об}, t2\_z1, \text{CHAR} \rangle, \langle \text{Об}, t2\_p2, \text{CHAR} \rangle, \langle \text{Об}, t2\_z2, \text{CHAR} \rangle, \langle \text{Об}, t2\_p3, \text{CHAR} \rangle,$

$\langle \text{Об}, t2\_z3, \text{CHAR} \rangle, \langle \text{Об}, t2\_gm, \text{CHAR} \rangle, \langle \text{Об}, up\_dt, \text{DATE} \rangle, \langle \text{Об}, up\_us, \text{CHAR} \rangle,$

$\langle \text{Об}, t2\_dc, \text{BLOB} \rangle, \langle \text{Об}, t2\_vid, \text{CHAR} \rangle, \langle \text{Об}, t2\_doc, \text{CHAR} \rangle,$

$\langle \text{Об}, t2\_prim, \text{CHAR} \rangle, \langle \text{Об}, t2\_yyyy, \text{CHAR} \rangle \}$

$R_L = \{ \langle t2\_ob, \langle \text{length}, 200 \rangle \rangle, \langle t2\_ng, \langle \text{precision}, 5 \rangle \rangle, \langle t2\_nn, \langle \text{precision}, 6 \rangle \rangle,$

$\langle t2\_p1, \langle \text{length}, 2 \rangle \rangle, \langle t2\_p1, \text{nullable} \rangle, \langle t2\_z1, \langle \text{length}, 8 \rangle \rangle, \langle t2\_z1, \text{nullable} \rangle,$

$\langle t2\_p2, \langle \text{length}, 2 \rangle \rangle, \langle t2\_p2, \text{nullable} \rangle, \langle t2\_z2,$

$\langle \text{length}, 8 \rangle \rangle, \langle t2\_z2, \text{nullable} \rangle, \langle t2\_p3, \langle \text{length}, 2 \rangle \rangle, \langle t2\_p3, \text{nullable} \rangle, \langle t2\_z3, \langle \text{length}, 8 \rangle \rangle, \langle t2\_z3, \text{nullable} \rangle, \langle up\_us, \langle \text{length}, 32 \rangle \rangle, \langle t2\_vid, \langle \text{length}, 4 \rangle \rangle, \langle t2\_doc, \langle \text{length}, 100 \rangle \rangle, \langle t2\_prim, \langle \text{length}, 100 \rangle \rangle, \langle t2\_doc, \text{nullable} \rangle, \langle t2\_yyyy, \langle \text{length}, 4 \rangle \rangle \}$

$\rangle$ .  
Онтологическое представление  $O^{receiver}$  сущности «Оборудование и инструменты» (таблицы 3, 4, 5, 6) системы приемника будет иметь следующий вид:

$O^{receiver} = \langle$

$C = \{ \text{Оборудование (Об), Тип оборудования (ТОб), Параметры оборудования (ПОб), Оборудование Параметры оборудования (Об-ПОб), INTEGER, VARCHAR, DATE},$

$P = \{ \text{Об\_id, Об\_name, serial\_number, inventory\_number, production\_date, tool\_types\_id, ТОб\_id, ТОб\_name, ПОб\_id, ПОб\_name, tool\_id, tool\_parameter\_id, value} \}$

$L = \{ \text{nullable, } \langle \text{length}, 255 \rangle \}$

$R_P = \{ \langle \text{Об}, \text{Об\_id}, \text{INTEGER} \rangle, \langle \text{Об}, \text{Об\_name}, \text{VARCHAR} \rangle,$

$\langle \text{Об}, \text{serial\_number}, \text{VARCHAR} \rangle, \langle \text{Об}, \text{inventory\_}$

```

number, VARCHAR>,
<Об, production_date, DATE>, <Об, tool_types_id,
INTEGER >,
<ТОб, ТОб_id, INTEGER>, <ТОб, ТОб_name,
VARCHAR>,
<ПОб, ПОб_id, INTEGER>, <ПОб, ПОб_name,
VARCHAR>,
<ОбПОб, tool_id, INTEGER>, <ОбПОб, tool_
parameter_id, INTEGER>,
<ОбПОб, value, VARCHAR>}
 $R_L = \{<Об\_name, <length, 255>>, <serial\_number,
<length, 255>>,
<serial\_number, nullable>, <inventory\_number,
<length, 255>>,
<ТОб\_name, <length, 255>>, <ПОб\_name, <length,
255>>, <value, <length, 255>>\}$ 
}

```

Для полученных онтологических представлений  $O^{source}$  и  $O^{receiver}$  процесс формирования интегрирующей модели данных  $O^{META}$  будет состоять из следующих шагов:

Шаг 1. Выделение общего понятийного аппарата ПрО:

1. Термин «Порядковый номер» системы источника соответствует термину «Первичный ключ» системы приемника.
2. Термин «Примечание» системы источника соответствует термину «Инвентарный номер» системы приемника.
3. Термин «Год производства» системы источника соответствует термину «Дата производства» системы приемника.

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

```

 $C^{META} = \{<Источник, Oracle, <login, password>>,
<Приемник, PostgreSQL, <login,
password>>\}$ 

```

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

```

 $R_C^{META} = \{<Источник, Оборудование и инстру-
менты>, <Приемник, Оборудование>,
<Приемник, Тип оборудования>,
<Приемник, Параметры оборудова-
ния>,
<Приемник, Оборудование Параме-
тры оборудования>\}$ 

```

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

```

 $R_P^{META} = \{<<Источник, t2\_nn>, t2\_nn\_Об\_id,
<Приемник, Об\_id>>,
<<Источник, t2\_об>, t2\_об\_Об\_name,
<Приемник, Об\_name>>,
<<Источник, t2\_prim>, t2\_prim\_inv\_
numbrt, <Приемник, inventory\_
number>>,
<<Источник, t2\_yyyy>, t2\_yyyy\_prod\_$ 

```

```

date, <Приемник, production_date>>,
<<Источник, t2\_r1>, t2\_r1\_ТОб\_name,
<Приемник, ТОб\_name >>,
<<Источник, t2\_p1>, t2\_p1\_ПОб\_name,
<Приемник, ПОб\_name>>,
<<Источник, t2\_z1>, t2\_z1\_value, <При-
емник, value>>,
<<Источник, t2\_p2>, t2\_p2\_ПОб\_name,
<Приемник, ПОб\_name>>,
<<Источник, t2\_z2>, t2\_z2\_value, <При-
емник, value>>,
<<Источник, t2\_p3>, t2\_p3\_ПОб\_name,
<Приемник, ПОб\_name>>,
<<Источник, t2\_z3>, t2\_z3\_value, <При-
емник, value>>\}

```

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

```

 $R_L^{META} = \{<t2\_об\_Об\_name, <Приемник, <length, 255>>,
<t2\_r1\_ТОб\_name, <Приемник, <length, 255>>,
<t2\_p1\_ПОб\_name, <Приемник, <length, 255>>,
<t2\_z1\_value, <Приемник, <length, 255>>,
<t2\_p2\_ПОб\_name, <Приемник, <length, 255>>,
<t2\_z2\_value, <Приемник, <length, 255>>,
<t2\_p3\_ПОб\_name, <Приемник, <length, 255>>,
<t2\_z3\_value, <Приемник, <length, 255>>\}$ 

```

После получения интегрирующей модели данных  $O^{META}$  запуск машины вывода  $M$  позволит проверить логическую целостность и непротиворечивость получившейся онтологии. Для проверки выполнения условий ограничений  $L^{META}$  требуется разработка соответствующих методов, так как существующие машины вывода не поддерживают данную функциональность.

Таким образом, в результате выполнения данных шагов на основе полученных онтологических представлений  $O^{source}$  и  $O^{receiver}$  моделей данных формируется интегрирующая модель данных  $O^{META}$ , позволяющая обеспечить полноту и непротиворечивость данных, а также минимизировать дополнительный ввод информации за счет направления потока данных от существующей на предприятии ИС в сторону системы оптимизации ресурсов авиастроительного предприятия.

## ЗАКЛЮЧЕНИЕ

В данной статье представлена реализация метода интеграции ИС авиастроительного предприятия с системой оптимизации ресурсов. 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.

Таким образом, были решены поставленные ранее методологические задачи:

1. Создана интегрирующая модели данных, являющейся основой единого пользовательского интерфейса в системе интеграции.

2. Разработан метод отображения ИО интегрируемых ИС в онтологическое представление модели данных.

3. Предложена методика слияния онтологических представлений модели данных ИО интегрируемых ИС в интегрирующую модель данных.

4. Произведена интеграция метаданных ИО интегрируемых ИС.

5. Решена проблема неоднородности источников данных.

6. Разработан механизм семантической интеграции источников данных.

В качестве направления дальнейших исследований можно выделить следующие задачи:

1. Разработка подхода к автоматическому обмену данными между ИС авиастроительного предприятия и системой оптимизации ресурсов на основе сформированной интегрирующей модели данных.

2. Разработка методов проверки выполнения условий ограничений, так как существующие машины вывода не поддерживают работу с такими объектами.

## СПИСОК ЛИТЕРАТУРЫ

1. Интеграция проектных диаграмм и онтологий в задаче балансировки мощностей авиастроительного предприятия / Н.Г. Ярушкина, Т.В., Негода

- В.Н. Афанасьева, М.К. Самохвалов, А.М. Наместников, А.А. Романов, Г.Ю. Гуськов // Автоматизация процессов управления. 2017. № 4. С. 85-95.
2. Коголовский М.Р. Методы интеграции данных в информационных системах // Институт проблем рынка РАН: [сайт]. URL: <http://www.ipr-ras.ru/articles/kogalov10-05.pdf> (дата обращения 16.10.2018).
3. Кусов А.А. Проблемы интеграции корпоративных информационных систем. URL: <https://cyberleninka.ru/article/n/problemy-integratsii-korporativnyh-informatsionnyh-sistem> (дата обращения 16.10.2018).
4. Морозова О.А. Интеграция корпоративных информационных систем: Учебное пособие. — М.: Финансовый университет, 2014. С. 8–23.
5. Степанов Д.Ю. Способы интеграции данных корпоративных информационных систем // Естественные и технические науки. — М., 2014. С. 207-213.
6. C. Bizer, T. Heath, T. Berners - Lee Linked Data – The Story So Far [Электронный ресурс] // Home – Tom Heath: [сайт]. URL: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf> (дата обращения 15.10.2018).
7. OWL 2 Web Ontology Language Document Overview (Second Edition) // World Wide Web Consortium: [сайт]. URL: <https://www.w3.org/TR/owl2-overview/> (дата обращения 15.10.2018).
8. T. Gruber Ontology // Tom Gruber [сайт]. URL: <http://tomgruber.org/writing/ontology-in-encyclopedia-of-dbs.pdf> (дата обращения 10.10.2018).
9. A. Poggi, D. Lembo, D. Calvanese, G. De Giacomo, M. Lenzerini, R. Rosati. Linking data to ontologies // Data Semantics, 2008, С.133–173.
10. D. Calvanese, B. Cogrel, S. Komla-Ebri, R. Kontchakov, D. Lanti, M. Rezk, M. Rodriguez-Muro, G. Xiao. Ontop: Answering SPARQL Queries over Relational Databases // Semantic Web journal: [сайт]. URL: <http://www.semantic-web-journal.net/system/files/swj1278.pdf> (дата обращения 11.10.2018).
11. Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем. – СПб.: Питер, 2000, С.59–98.

## INTEGRATION OF INFORMATION SYSTEMS BASED ON ONTOLOGY MERGING FOR THE BALANCING OF INDUSTRIAL PROCESSES

© 2018 N.G. Yarshkina, A.A. Romanov, A.A. Filippov, A.Y. Dolganovskaya, M.S. Grigoricheva

Ulyanovsk State Technical University

This paper presents an approach to data integration based on an integrating model. An ontological data model is created, a data mapping from information systems to an ontological model is made. An illustrative example is considered.

*Keywords:* ontology, integration of information systems, data model.

*Nadezhda Yarushkina, Doctor of Technics, Professor at the Information Systems Department. E-mail: jng@ulstu.ru*

*Anton Romanov, Candidate of Technics, Associate Professor at the Information Systems Department.*

*E-mail: romanov73@gmail.com*

*Aleksey Filippov, Candidate of Technics, Associate Professor at the Information Systems Department.*

*E-mail: al.al.filippov@gmail.com*

*Aleksandra Dolganovskaya, Postgraduate Student.*

*E-mail: adolganovskaya@mail.ru*

*Mariya Grigoricheva, Postgraduate Student.*

*E-mail: gms4295@mail.ru*