

Федеральное государственное автономное образовательное учреждение
высшего образования «Уральский федеральный университет
имени первого Президента России Б.Н. Ельцина»

На правах рукописи



Пономарева Ольга Алексеевна

РАЗРАБОТКА МЕТОДОЛОГИИ СТРУКТУРНОГО СИНТЕЗА ХРАНИЛИЩ
ГЕТЕРОГЕННЫХ ДАННЫХ ПРОМЫШЛЕННОГО ПРЕДПРИЯТИЯ

2.3.1. Системный анализ, управление и обработка информации

АВТОРЕФЕРАТ
диссертации на соискание ученой степени кандидата
технических наук

Екатеринбург – 2021

Работа выполнена в учебно-научном центре «Информационная безопасность» Института радиоэлектроники и информационных технологий-РтФ ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина».

Научный руководитель

доктор технических наук, профессор,
Поршнев Сергей Владимирович

Официальные оппоненты

Зобнин Борис Борисович, доктор технических наук, профессор, ФГБОУ ВО «Уральский государственный горный университет», г. Екатеринбург, профессор кафедры информатики;

Лавров Вячеслав Васильевич, доктор технических наук, доцент, ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина», профессор кафедры теплофизики и информатики в металлургии;

Успенский Михаил Борисович, кандидат технических наук, ФГАОУ ВО «Санкт-Петербургский политехнический университет Петра Великого», г. Санкт-Петербург, научный сотрудник лаборатории «Промышленные системы потоковой обработки данных» научного центра мирового уровня «Передовые цифровые технологии»

Защита диссертации состоится «07» декабря 2021 г. в 15:00 ч на заседании диссертационного совета УрФУ 2.3.05.13 по адресу: 620002, г. Екатеринбург, ул. Мира, 19 ауд. И-420 (зал Ученого Совета).

С диссертацией можно ознакомиться в библиотеке и сайте ФГАОУ ВО «Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»:

<https://dissovet2.urfu.ru/mod/data/view.php?id=12&rid=2892>

Автореферат разослан «__» _____ 2021 г.

Ученый секретарь
диссертационного совета



Сафиуллин Николай Тахирович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. В современных условиях для управления бизнес-процессами каждого современного промышленного производства используют различные информационные системы (ИС), которые призваны решать, как задачи управления технологическим оборудованием, используемого для производства продукции, так и задачи сбора, хранения и анализа разнородной информации. Как правило, промышленные ИС по укрупненной классификации (Е.П. Стефани) разделяются на: АСУ ТП (автоматизированные системы управления технологическими процессами), MES-системы (Manufacturing Execution System) – автоматизированные системы управления производственной деятельностью предприятия в режиме реального времени, ERP-системы (Enterprise Resource Planning), обеспечивающие автоматизацию планирования, учета, контроля и анализа каждого бизнес-процесса промышленного предприятия. При этом оказывается, что каждая из данных типов ИС имеет собственные, уникальные источники информации, а также скорость генерации информации и структуру информации.

В результате информация о жизненном цикле одной и той же единицы продукции (ЕП) промышленного производства, которая необходима для решения задач управления, оптимизации бизнес-процессов, выявления и устранения причин брака, оказывается размещенной в нескольких автономных хранилищах данных (ХД).

С информационной точки зрения сущность «единица готовой продукции (ГП)» представляет собой кортеж

$$ГП = \left\{ \left\{ П^{АСУТП} \right\}, \left\{ П^{MES} \right\}, \left\{ П^{ERP} \right\}, t \right\}, \quad (1)$$

где

$$\begin{aligned} \left\{ П^{АСУТП} \right\} &= \bigcup_{i=1, N_{АСУТП}} \left\{ \left\{ П_i^{АСУТП} \right\}, t \right\}, \\ \left\{ П^{MES} \right\} &= \bigcup_{j=1, N_{MES}} \left\{ \left\{ П_j^{MES} \right\}, t \right\}, \\ \left\{ П^{ERP} \right\} &= \left\{ \left\{ Договор \right\}, \left\{ Заказ \right\}, \left\{ План \right\}, \left\{ ЕП \right\}, t \right\}, \end{aligned}$$

где

$\left\{ П_i^{АСУТП} \right\}$ – совокупность параметров технологических процессов, собираемых i -ой АСУ ТП;

$\left\{ П_j^{MES} \right\}$ – совокупность параметров бизнес-процессов и производственных процессов, собираемых j -ой MES-системой;

$\left\{ Договор \right\}$ – совокупность параметров бизнес-процессов, относящихся к заключенным договорам о производстве ГП;

$\left\{ Заказ \right\}$ – совокупность параметров бизнес-процессов, относящихся к заказам, выданным производственным службам, на основании заключенных договоров о производстве ГП;

$\left\{ План \right\}$ – совокупность параметров плана производства ЕП, сформированного на основании заключенных договоров и заказов на производство продукции;

$\left\{ ЕП \right\}$ – совокупность параметров ЕП, характеризующих состояние ЕП на каждом из этапов производственного процесса, а также, собственно, параметров технологических процессов.

С точки зрения распределения данных, относящихся к сущности «ГП», эта сущность интегрирует в себе множество разнородных данных, которые собираются в автономные БД АСУ ТП, MES- и ERP-систем. Как следствие, возникает проблема поиска информации, релевантной решаемой задаче, которая находится в нескольких автономных ХД. При этом дополнительные трудности возникают в связи с тем, что в большинстве случаев информация,

хранящаяся в ХД одной ИС, оказывается не синхронизованной во времени с информацией, хранящейся в ХД в другой ИС.

В полной мере перечисленные выше проблемы, связанные с вынужденным распределением технологической информации по нескольким автономным ХД, и необходимости организации в этой связи информационного взаимодействия между различными ИС, присущи современному металлургическому производству (МП), на котором реализуется полный жизненный цикл продукции: руда-чугун-сталь-металлопрокат, а также, при необходимости, утилизация продукции МП, что, закономерно, приводит к динамическому изменению структуры технологической информации, соответствующей ее текущему состоянию, на каждом из этапов жизненного цикла (ЖЦ) продукции МП.

Выход из сложившейся ситуации достаточно очевиден – перейти от использования автономных ХД к использованию единого ХД, пригодное для хранения всей информации, создающейся в процессе производства продукции промышленного предприятия, что позволит хранить, извлекать и анализировать информацию, структура которой принципиально разнородна и может изменяться с течением времени.

Для создания единого ХД промышленного производства требуется решить многочисленные проблемы интеграции разнородных и, зачастую, несогласованных друг с другом, создаваемых разнородными источниками, работающими на различных физических принципах и обеспечивающим различный темпы генерации информации, которые далее мы будем называть гетерогенными данными (ГД), соответственно единое ХД – хранилищем гетерогенных данных (ХГД).

Степень проработанности темы. Анализ работ, посвященных различным аспектам разработки ХД и проблемам интеграции разнородных данных, показал, что сегодня создана теоретическая база для разработки моделей данных, в том числе:

– реляционные модели данных (Е. F. Codd, S. B. Codd and C. T. Salley, В.И. Халимон, Г.А. Мамаева, А.Ю. Рогов, В.Н. Чепикова; В.И. Швецов, А.Н. Визгунов, И.Б. Мееров, Ю.П. Парфенов, Т.С. Карпова и др.);

– объектно-ориентированные модели данных (Г. Буч, Beunon-Davies, А. В. Сиха Багуи Зафиевский, А. А. Короткин, А. Н. Лататуев, В.И. Швецов, А.Н. Визгунов, И.Б. Мееров, И.А. Микляев и др.);

– темпоральные модели данных (А.В. Балдин, Д.В. Елисеев, К.Г. Агаян, Д.В. Елисеев и др.).

Существуют технологии, предназначенные для создания ХД различных типов, в том числе:

– технология транзакционной организации записи OLTP (Online Transaction Processing); предназначенная для ХД с заранее выбранной и далее не модифицируемой структурой хранения данных (А. В. Зафиевский, А. А. Короткин, А. Н. Лататуев, В.И.Швецов, А.Н. Визгунов, И.Б. Мееров, Ю.П. Парфенов, Т.С. Карпова, И. Афанасьев и др.);

– технология OLAP (Online Analytical Processing), ориентированная на хранение большого объема структурированных данных, относящихся к определенной предметной области (Э. Спирли, Сураджит Чаудхури, Умешвар Дайал, Венкатеш Ганти, С. Федечкин, С. Архипенков, Д. Голубев, О. Максименко; А.М. Бородин, А. Барсегян, М. С. Куприянов, В. В. Степаненко, И. И. Холод., А.Б. Бергер и др.);

– технология Oracle Streams (см. <https://oracle-patches.com/oracle/prof>), предназначенная для интеграции данных, обмена данными и сообщениями с помощью механизма Advanced Queuing (см. <https://www.oracle.com/database/technologies/advanced-queuing.html>) в однородной среде и гетерогенных средах.

Также были поставлены задачи интеграции разнородных данных и разработаны подходы к их решению (David Taniar, Li Chen, D. Beneventano, S. Bergamaschi, В.Е. Prasad, P.G. Reddy, Amar Gupta, Д. Бин, Г. Хоп, Б. Вульф, О.А. Морозова, Н. Игнатович Флориан Дениэль, Мариселла Матера, Ю. Джин, Булем Бенаталлах, Реджис Сен-Поль, Фабио Касати, Л. А. Калиниченко, И. А. Думченков, Л. Черняк, К. А. Аксенов и др.), которые, однако, не являются универсальными.

Также отметим, что сегодня в соответствующем сегменте рынка программного обеспечения представлено большое число систем управления базами данных (СУБД), например, Oracle (<https://www.oracle.com/index.html>), MS SQL (<https://www.microsoft.com/ru-ru/sql-server/sql-server-downloads>) и др.

Однако, несмотря на существенный прогресс, достигнутый в теории проектирования ХД и разработки методов интеграции неоднородных, по сути, гетерогенных данных (ГД), как показали результаты анализа опыта применения ИС и ХД на МП, на момент начала исследования не существовало ни одного единого ХГД МП.

Данная ситуация, с нашей точки зрения, обусловлена, в первую очередь, отсутствием методологии структурного синтеза единого ХГД, под которой мы понимаем совокупность моделей (структурные модели информационных потоков промышленного производства; информационные и онтологические модели источников информации) и технологии разработки ХГД, выбор которых должен базироваться на результатах системного анализа выбранного промышленного производства и соответствующих информационных потоков.

Основной целью исследования является разработка методологии структурного синтеза ХГД и ее апробация на примере МП.

Для достижения цели исследования требуется решить следующие **задачи**:

1. Проанализировать методы интеграции разнородных данных, технологий хранения и доступа к разнородным данным, их современное состояние, условия создания единого хранилища разнородных данных.

2. Обосновать гетерогенную природу данных производственных процессов многоэтапного промышленного предприятия, анализ которых позволит решать задачи управления качеством продукции.

3. Разработать универсальную методологию структурного синтеза единого ХГД, провести системный анализ технологических процессов, обосновать применение информационных сущностей «План производства», «Готовая продукция», «Единица продукции», привести их онтологическое описание, разработать логическую структуру единого ХГД и обосновать выбор технологии создания единого ХГД, которая обеспечит реализацию логической структуры ХГД.

4. Разработать и провести тестирование модуля ХГД МП, как части Автоматизированной системы выпуска металлургической продукции (АС ВМП).

Объект исследования – методы интеграции ГД, создаваемых в процессе промышленного производства.

Предмет исследования – Информационные и онтологические модели источников данных МП.

Научная новизна:

1. Обоснована сущность ГП МП, которая состоит из двух сущностей <ПП, ЕП>.

2. Разработана структурная схема сущностей ГП, ПП, ЕП, а также информационная и онтологическая модели источников данных, что обеспечивает возможность структурного синтеза ХГД МП.

3. Разработана научно обоснованная методология структурного синтеза ХГД промышленных предприятий.

Теоретическая значимость диссертационного исследования заключается в: разработке методологии структурного синтеза единого ХГД МП, включающей в себя структурные модели информационных потоков МП; информационные и онтологические модели источников информации МП; синтезированную логическую структуру ХГД МП и обоснованно выбранную технологию разработки ХГД, которую можно тиражировать на других промышленных предприятиях.

Практическая значимость диссертационного исследования заключается в: разработке на основе созданной методологии синтеза ХГД МП программного модуля «ХГД МП», как составной части системы выпуска металлургической продукции, и подтверждение его работоспособности результатами тестовых испытаний.

Результаты диссертационного исследования были получены и использованы в ходе выполнения договора № 02.G25.31.0055 (проект 2012-218-03-167) на разработку АС ВМП.

Методология и методы исследования. В работе использованы методы системного анализа и структурного синтеза, информационного и онтологического моделирования.

Основные положения, выносимые на защиту:

1. Разработанная информационная модель ГП МП позволяет решить задачу структурного синтеза ХГД МП.

2. Разработанная методология структурного синтеза ХГД МП является универсальной и может быть использована на других промышленных предприятиях.

3. Разработанные программные инструменты управления ХГД МП, обеспечивают его функционирование в составе АСВМП и работоспособность системы в целом

Достоверность результатов исследования подтверждается использованием математических методов, адекватным задачам исследования, а также результатами тестовых испытаний модуля «ХГД МП».

Внедрение результатов диссертационного исследования. Результаты диссертационного исследования используются в Уральском федеральном университете имени первого Президента России Б.Н. Ельцина, Екатеринбург, Россия (акт об использовании результатов от 20.08.2021); в АО «Ай-Тек», Москва, Россия (акт об использовании результатов от 01.09.2021); в ООО «Октоника», Екатеринбург, Россия (акт об использовании результатов от 10.09.2021).

Апробация работы. Результаты работы докладывались и обсуждались на научно-практических конференциях: Международные научно-технические конференции «Динамика Систем, Механизмов и Машин» (Омск, 2016, 2017); Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology – USBEREIT (Екатеринбург, 2018-2020); Новые Информационные Технологии в Исследовании Сложных Структур (Томск, 2018); International Conference on Applied Mathematics and Computational Science - ICAMCS (Budapest, Hungary, 2018); International Multi-Conference on Industrial Engineering and Modern Technologies - FarEastCon (Владивосток, 2019); Межвузовская конференция «Задачи Системного Анализа, Управления и Обработки Информации» (Москва, онлайн, 26 октября 2020), III международная студенческая научная конференция «Инновационные механизмы управления цифровой и региональной экономикой» (Москва, онлайн, 17-18 июня 2021).

Диссертационное исследование выполнено при финансовой поддержке гранта РФФИ № 19-37-90049 «Разработка методов интеграции и доступа к гетерогенным данным металлургического производства».

Личный вклад автора состоит в проведении системного анализ информационных потоков МП, разработке информационных и онтологических моделей ГД МП, логической структуры ХГД МП, методологии структурного синтеза ХГД, участие в разработке методики тестирования модуля ХГД МП и его проведении.

Публикации. По теме диссертации опубликовано 12 научных работ, в том числе 11 статей опубликованы в рецензируемых научных журналах и изданиях, определенных ВАК РФ и Аттестационным советом УрФУ, включая 9 статей в изданиях, индексируемых в международных цитатно-аналитических базах Scopus и Web of Science.

Структура и объем диссертации Диссертация состоит из введения, четырех глав, заключения, списка литературы из 102 наименований, содержит 57 рисунка и 24 таблиц. Основной текст работы составляет 131 страницу, приложения составляют 49 страниц. Общий объем – 180 страниц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении изложена общая характеристика работы, обоснована ее актуальность, сформулирована цель работы, описаны задачи, отражена научная новизна и практическая ценность.

В первой главе проведен анализ состояния предметной области, в том числе: методов интеграции разнородных данных, моделей разнородных данных и технологий хранения данных; определены следующие базовые понятия, используемые в работе: *гетерогенные данные* (результат интеграции данных, описываемых различными информационными моделями); *источники гетерогенных данных* (источники данных, описываемых различными информационными моделями); *хранилище гетерогенных данных* (хранилище данных, предназначенное для сбора, хранения и доступа к гетерогенным данным); *гетерогенная информационная система* (ИС, обеспечивающая сбор, анализ и обработку гетерогенных данных); *гетерогенный запрос* (запрос, обеспечивающий извлечение информации из гетерогенного хранилища данных). Сформулированы цель и задачи исследования.

Во второй главе проведен системный анализ информационных потоков (ИП) МП, обобщенная схема которых представлена на рисунке 1 и разработана структурная схема источников информации (ИИ) (рисунок 2).

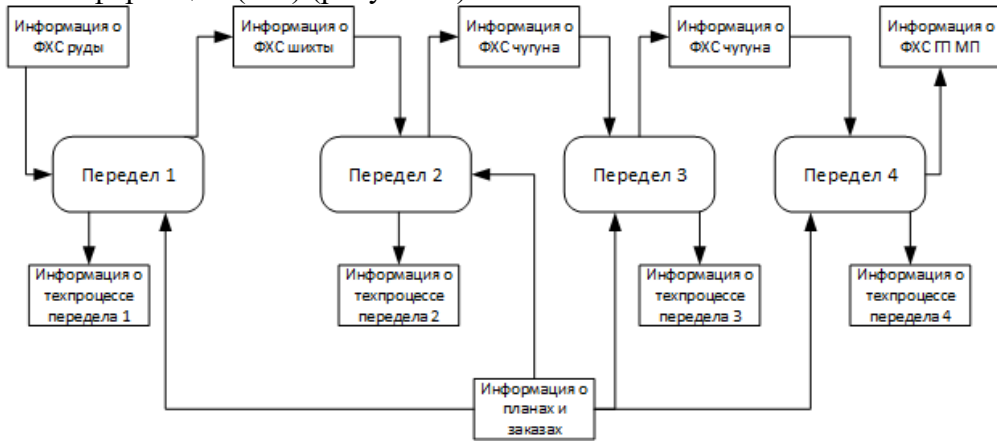


Рисунок 1. Обобщенная схема информационных потоков данных МП

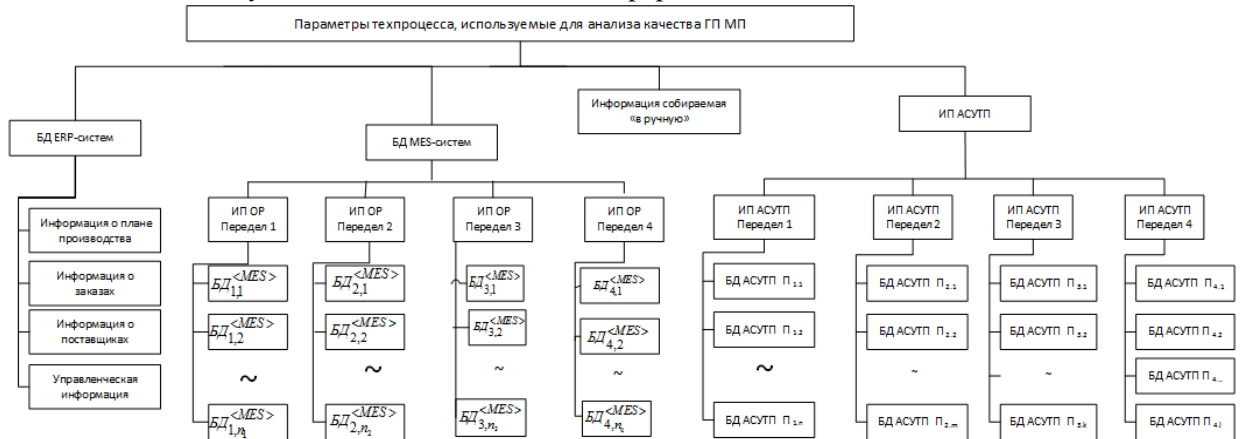


Рисунок 2. Структурная схема ИИ МП

На основе результатов проведенного системного анализа ИП сделан обоснованный вывод о том, что данные МП можно отнести к одному из трех типов: статистические данные, квазидинамические данные и динамические данные. Для подтверждения адекватности предложенной классификации проведен анализ информационных параметров этапа непрерывного литья заготовки (НЛЗ) на стане 2000, результаты которого представлены в таблице 1.

Таблица 1. Количественная оценка информационных параметров

Параметр	Количество статических параметров	Количество квазидинамических параметров	Количество динамических параметров	Общее число контролируемых параметров
Перечень регистрируемых параметров прокатки и слежения за металлом на линии стана 2000	10	56	97	163
Параметры сляба	23	39	22	84

Из таблицы 1 видно, что общее число контролируемых параметров машины НЛЗ оказывается равным 163 (10 статических параметров, 56 квазидинамических параметров, 97 динамических параметров), общее число контролируемых параметров сляба – 84 (23 статических параметра, 39 квазидинамических параметра, 33 динамических параметра), что подтверждает гетерогенную природу данных МП (ровно, как и любого другого этапа производства промышленной продукции). В этой связи понятно, что для решения задач анализа причин брака продукции МП и повышения ее качества необходимо перейти от использования автономных БД ERP, различных БД MES и БД АСУ ТП, к хранению информации в едином ХГД МП.

Третья глава посвящена разработке методологии создания единого ХГД МП, в ходе которой созданы информационные модели следующих понятий: «план производства» (ПП), «готовая продукция» (ГП), «единица продукции» (ЕП), представляющие собой кортежи:

$$ПП_m = \left\{ \text{Номер ПП}_m, \overline{\PhiХС}_{j_m}^{(m)}, V^{(m)}, t^{(m)} \right\}, \quad (2)$$

где

$m = \overline{1, 4}$ – номер передела МП,

Номер $ПП_m$ – идентификатор ПП выпуска $ЕП_m$,

$\overline{\PhiХС}_{j_m}^{(m)}$ – $j_m = \overline{1, J_m}$, – множество ФХС $ЕП_m$, $J_{m-1} = \left| \overline{\PhiХС}^{(m-1)} \right|$,

$V^{(m)}$ – объем (в выбранных единицах измерения) выпуска $ЕП_m$,

$t^{(m)} = \left[t_{start}^{(m)}, t_{finish}^{(m)} \right]$ – временной интервал, в течение которого был реализован выпуск $ЕП_m$.

Соответственно, ЕП МП есть совокупность информации, характеризующей ЕП на каждом из этапов МП:

$$ЕП = \{ ЕП_1, ЕП_2, ЕП_3, ЕП_4 \}, \quad (3)$$

здесь

$$ЕП_m = \{ \text{Номер ПП}_m, \PhiХС_{i_m}^{(m)}, ПТП_{k_m}^{(m)}, \overline{\PhiХС}_{j_m}^{(m)}, V^{(m)}, t^{(m)} \}, \quad (4)$$

где

$\PhiХС_{i_m}^{(m)}$, $i_1 = \overline{1, I_1}$ – множество значений ФХС компонентов, используемых для производства $ЕП_m$, $I_m = \left| \PhiХС^{(m)} \right|$,

$ПТП_{k_m}^{(m)}$, $k_m = \overline{1, K_m}$ – множество значений параметров технологических процессов (ПТП), реализованных для производства $ЕП_m$, $K_m = \left| ПТП^{(m)} \right|$,

$\overline{\PhiХС}_{j_m}^{(m)}$, $j_m = \overline{1, J_m}$, – множество ФХС $ЕП_m$, $J_m = \left| \overline{\PhiХС}^{(m)} \right|$,

$V^{(m)}$ – объем (в выбранных единицах измерения) выпуска $ЕП_m$,

$t^{(m)} = \left[t_{start}^{(m)}, t_{finish}^{(m)} \right]$ – временной интервал, в течение которого был реализован выпуск $ЕП_m$.

Из (2)–(4) видно, что информация о ПП, ГП МП и ЕП МП, является гетерогенной информацией. Она, вообще говоря, имеет четырехслойную структуру.

В связи с тем, что i -ом переделе МП причиной реализации соответствующего технологического процесса является некоторое событие C_i , возникающие в процессе выпуска EP_i (например, поступление EP_{i-1} с предыдущего передела, возврат продукции EP_i в случае обнаружения брака на один из предыдущих переделов и т.д.), сделан обоснованный вывод о том, что структура единого ХГД, наряду с описанными выше информационными структурами, должна обеспечивать хранение информации, структура которой есть триплет «событие, ЕП, источник информации»:

$$ДЕП_i = \langle C_i, EP_i, ИИ_i \rangle, \quad (5)$$

который описывает «динамику» ЕП.

Затем были разработаны онтологические модели ПП, ГП, ЕП МП, на основе которых далее описаны структурные взаимосвязи сущностей ЕП МП, включающие в себя: структурные схемы сущностей «ГП МП» (рис.3), «ЕП» (рис.4).

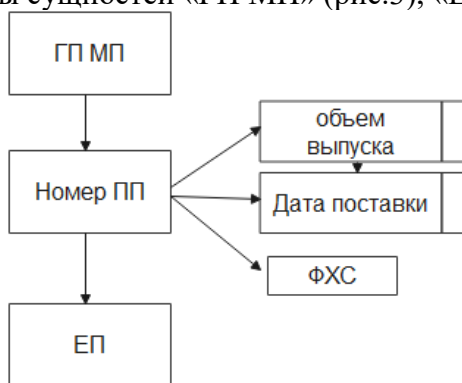


Рисунок 3. Структурная схема сущности «ГП МП»

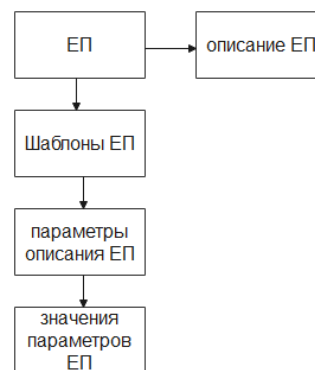


Рисунок 4. Структурная схема сущности «ЕП»

Из рисунка 4 видно, что сущность «ЕП», в которой на уровне атрибутов хранится описание ЕП, связана с сущностями «шаблон ЕП», «параметры описания ЕП», значения параметров ЕП». Здесь сущность «шаблон ЕП» позволяет создавать и хранить макеты описания ЕП для различных производственных переделов; сущность «параметры описания ЕП» – на каждом переделе задавать собственные шаблон и хранить регламентированные спецификацией данный ЕП параметры; сущность «значения параметров ЕП» обеспечивает хранение значение параметров согласно спецификации, данной ЕП.

Структурная схема сущностей, используемых для формирования структуры ЕП представлена на рисунке 5.

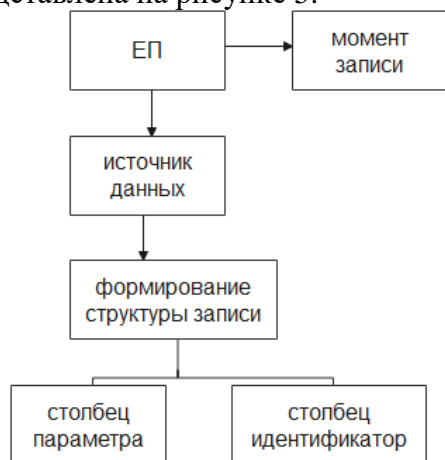


Рисунок 5. Структурная схема записи сущности «ЕП» в единое ХГД МП

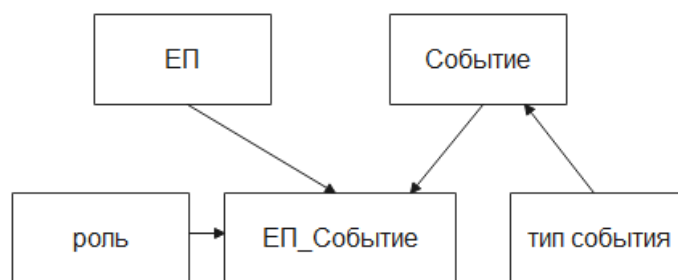


Рисунок 6. Схема связей сущностей, используемых для описания процесса хранения данных о событиях, возникающих в процессе жизненного цикла ЕП

Из рисунка 5 видно, что процесс формирования структуры ЕП описывается с помощью следующих сущностей: сущность «ЕП», сущность «момент записи», сущность «источник данных», сущность «формирование структуры записи», сущность «столбец параметра» и сущность «столбец идентификатор».

Схема сущностей, используемых для описания процесса хранения информации о событиях, происходящих в течение жизненного цикла ЕП МП, представлена на рисунке 6, из которого видно, что схема связей сущностей, используемых для описания процесса хранения данных о событиях, возникающих в процессе жизненного цикла ЕП, описывается помощью следующих сущностей: сущность «событие», сущность «роль», сущность «тип события». Сущность «событие» обеспечивает создание и хранение контрольных точек производственного процесса, в которых создается фиксируемая информация о ЕП МП (например, начало производства партии продукции, завершение производства партии продукции, отклонение параметром производственного процесса и другие факты).

Схема сущностей, используемых для описания процесса хранения информации о событиях, происходящих в течение жизненного цикла ЕП МП, представлена на рисунке 7.

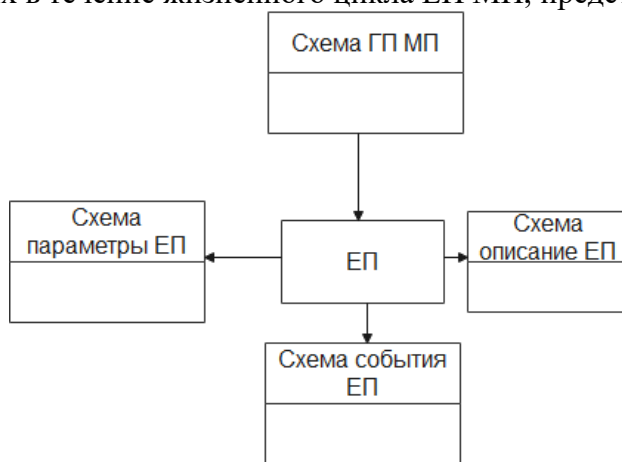


Рисунок 7. Схема связей сущностей, используемых для описания схем атрибутов ЕП

На основе онтологических моделей и схем сущностей, представленных на рисунках 5–7, разработана схема взаимодействия информационных потоков, взаимосвязанных с процессом выпуска ГП МП (рисунок 8), и далее логическая структура единого ХГД МП (рисунок 9). Данный результата позволяет сделать обоснованный вывод о том, что тем самым, решена задача структурного синтеза ХГД МП.

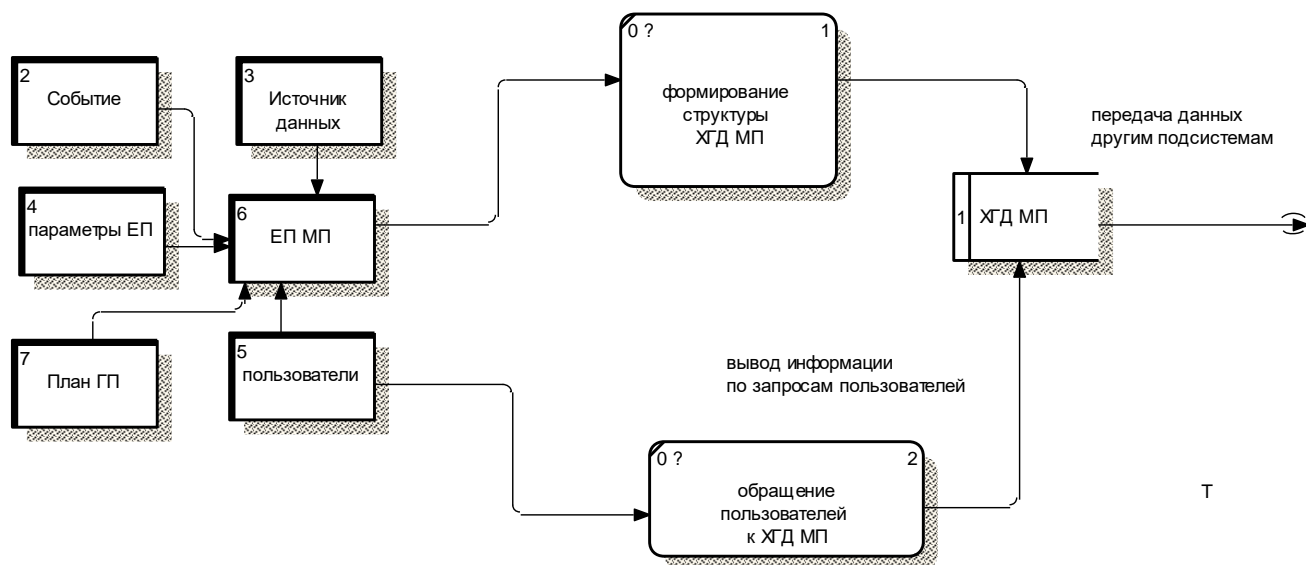


Рисунок 8. Диаграмма потоков данных, размещаемых в едином ХГД МП

Из рисунков 8, 9 видна основная особенность единого ХГД МП, состоящая в том, что структура записи данных в единое ХГД МП формируется в момент получения данных от соответствующих источников информации. В этой связи, в едином ХГД МП одновременно с данными необходимо размещать информацию об их структуре, на основе которой далее будет выполняться поиск информации, релевантной запросу пользователя к единому ХГД МП.

Анализ структурной схемы единого ХГД МП позволил обосновать следующие требования к технологии, используемой для создания единого ХГД МП: формирование структуры хранения данных в момент записи массива информации; выполнение транзакции на запись данных без операции отката; денормализацию данных с целью исключения множественных зависимостей от первичных ключей; индексацию/идентификацию данных в процессе формирования структуры данных; определение пользователей и их прав на этапе формирования структуры единого ХГД МП; обеспечение хранения данных, поступавших на протяжении всего жизненного цикла ГП МП; установление длительности хранения информации хранения информации в единое ХГД МП; запись в единое ХГД МП первичных данных, не прошедших процедуры агрегации и очистки; аналитический анализ всех бизнес-процессов МП.

Анализ существующих технологий хранения данных показал, что наиболее подходящей является технология Oracle Streams, предназначенная для интеграции данных, обмена данными и сообщениями с помощью механизма Advanced Queuing, в однородной среде и гетерогенных средах.

Таким образом, найдено решение задачи структурного синтеза ХГД МП, состоящее из следующих этапов.

Этап 1. Построение на основе анализа производственных процессов информационных моделей ПП, ГП, ЕП выбранного производства.

Этап 2. Построение на основе анализа информационных моделей ПП, ГП, ЕП выбранного производства онтологии гетерогенных данных выбранного производства инфологических моделей ПП, ГП, ЕП.

Этап 3. Разработка в соответствие со стандартом ГОСТ Р ИСО 10 303–11–2009 инфологической модели ХГД МП производства, основанной на инфологических моделях ЕП, ПП, ГП.

Этап 4. Выбор СУБД разработки единого ХГД МП.

Этап 5. Разработка и тестирование единого ХГД МП.

Описанная выше последовательность решения задачи структурного синтеза ХГД МП, представляющая собой совокупность научно обоснованных принципов, способов проведения системного анализа и моделирования информационных потоков МП, можно рассматривать как методологию структурного синтеза ХГД МП. При этом данная методология применима для каждого промышленного предприятия, в котором реализован поэтапный процесс выпуска продукции, а потому она является универсальной.

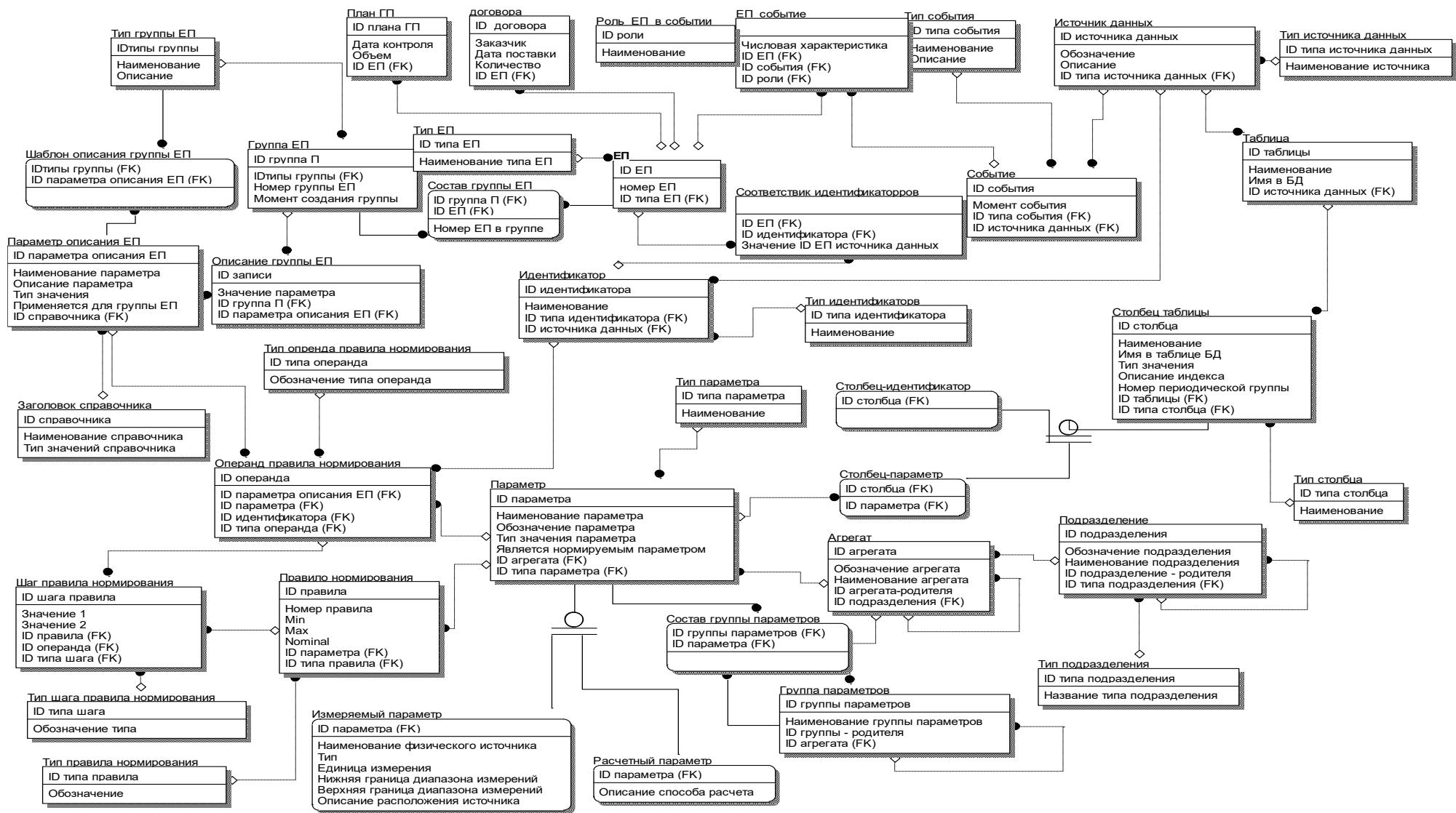


Рисунок 9. Логическая структура ХГД МП

В четвертой главе на основании предложенной методологии структурного синтеза единого ХГД МП разработаны программные инструменты, обеспечивающие функционирование модуля ХГД МП, являющегося неотъемлемой частью Автоматизированной Системы Выпуска Metallургической Продукции (АС ВМП), разработанной в рамках договора № 02.G25.31.0055 (проект 2012-218-03-167). Структурная схема АС ВМП представлена на рисунке 10.

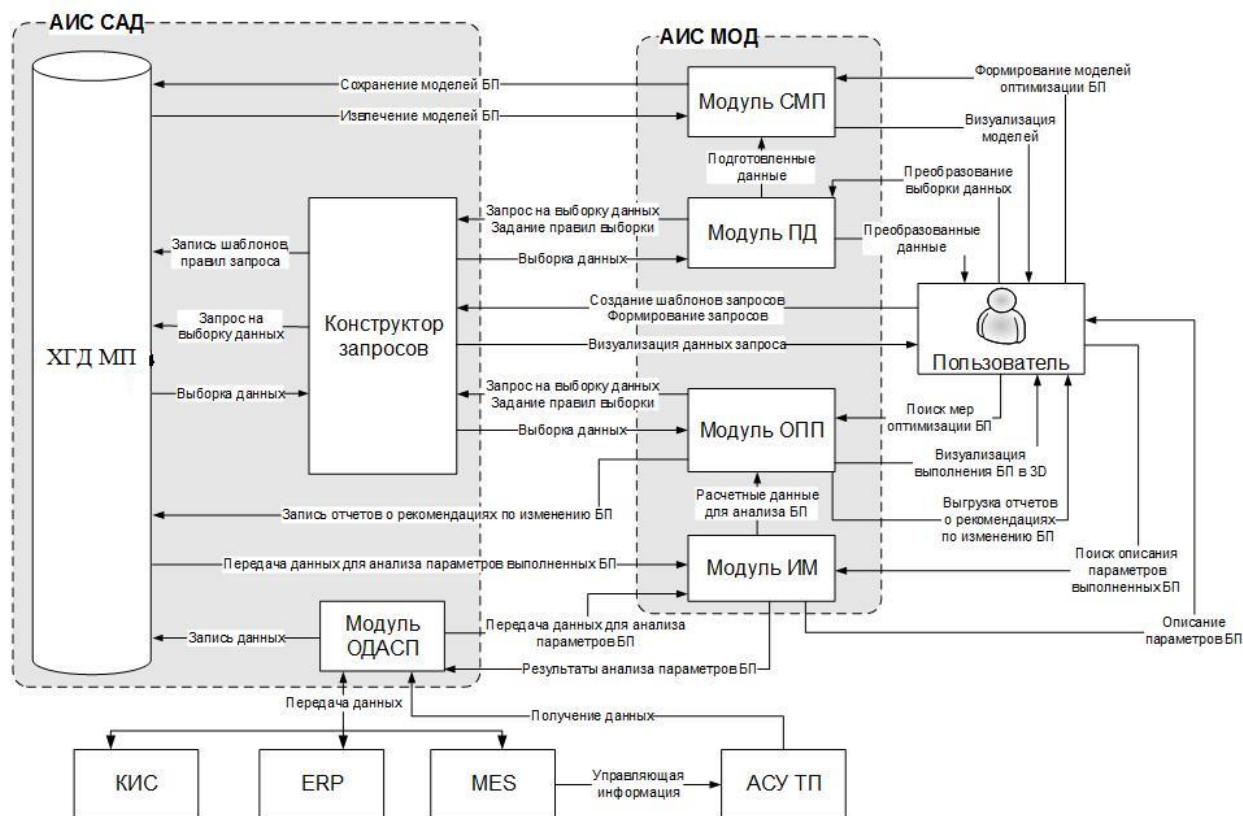


Рисунок 10. Структура АС ВМП: АИС САД – Автоматизированная Информационная Система Сбора и Анализа Данных; АИС МОД – Автоматизированная Информационная Система Моделирования Организационной Деятельности предприятия; ОДАСП – Обмена Данными с Автоматизированными Системами Предприятия; КИС – Корпоративные информационные системы; ОПП – модуль Организационных Процессов Предприятия; СМП – Модуль Создания Моделей Предприятия; ПД – модуль Подготовки Данных; ИМ – Модуль Имитационного Моделирования.

Разработанные программные инструменты обеспечили решение следующих задач:

- сбор, хранение и предоставление технологической информации по запросам пользователей; надежное хранение и быстрый доступ к ХГД МП, поддержка их хронологии, целостности и непротиворечивости, что обеспечивает осуществление высокопроизводительной аналитической обработки данных;
- восстановление данных; обмен данными между ХГД МП АИС САД и автоматизированными системами МП всех уровней ОДАСП;
- предоставление пользователю удобного интерфейса для создания требуемых выборок информационных параметров и формирование нерегламентированных отчетов (КЗ); предоставление данных для создания имитационных моделей технологических, логистических и организационных (бизнес) процессов (СМП).

Проведенное опытное тестирование ХГД МП выявило необходимость оптимизации выполнения запросов к ХГД МП. В этой связи было предложено совместно использовать: индекс MD (многомерный индекс на основе R-дерева, обеспечивающий поиск по множеству параметров ЕП, что соответствует структуре хранения данных в ХГД МП); индекс ET (битовый индекс выполнения условия, обеспечивающий поиск точного соответствия

условию); индекс FE (индекс, кэширующий пересчитанные значения выражений value в хэш-таблице), и создан соответствующий программный инструмент, в котором реализовано данное решение. Структурная схема механизма выполнения запросов на языке КЗ и команд ХГД МП с помощью кэша данных приведена на рисунке 11.

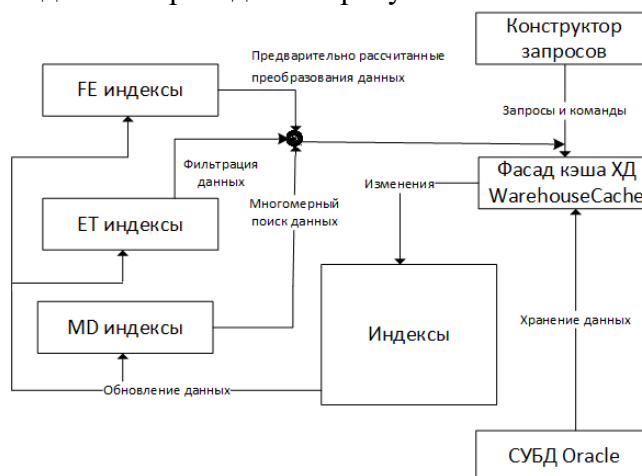


Рисунок 11. Структурная схема взаимодействия КЗ с ХГД МП при использовании кэширования данных

Из рисунка 11 видно, что кэш данных ХГД МП, размещаемый в основной памяти ХГД МП, обеспечивает осуществляет вызовы данных, соответствующих условиям запроса, из ХГД МП в кэш основной памяти без использования вызовов СУБД Oracle. При этом кэш данных настраивается через шаблонные запросы конструктора запросов, в котором в опциях колонки можно указать какие данные и как должны быть проиндексированы.

Для подтверждения работоспособности предложенных решений в диссертационном исследовании разработана методика проведения испытаний, предусматривающая проверку выполнения следующих функций:

1. формирования массива данных по заданной структуре;
2. создания и запись шаблонов построенных запросов в ХГД МП;
3. проверки использования технологии кэширования данных ХГД в основной памяти и синхронизация кэша данных с ХГД МП;
4. выполнения запросов на выдачу данных о переназначении продукции между заказами на поставку и отбраковку продукции;
5. выполнения запросов на выдачу данных о параметрах выпущенной продукции;
6. выполнения запросов на выдачу данных о функционировании и простоях оборудования;
7. выполнения запросов на выдачу данных о перемещениях продукции по переделам в ходе технологического процесса.

Результаты тестовых испытаний, разработанных программных инструментов, обеспечивающих функционирование модуля ХГД МП в составе единой АС ВМП, приведенные в разделе 4.5 диссертации, подтвердили их работоспособность. Протоколы испытаний, полные листинги разработанного программного обеспечения могут быть предоставлены по запросу на электронную почту o.a.ponomareva@urfu.ru.

В заключении подведены итоги исследования и сформулированы основные научные и практические результаты.

В приложениях приведены характеристики информационных параметров технологических процессов в металлургическом производстве, листинги разработанных программных инструментов, обеспечивающих функционирование ХГД МП, акты внедрения результатов диссертационного исследования.

ОСНОВНЫЕ ВЫВОДЫ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

1. Проведен анализ проблем интеграции разнородных данных, методов интеграции, технологий хранения и доступа к разнородным данным показал, что их современное состояние потенциально позволяет создавать хранилища разнородных данных, однако, для этого необходима разработка методологии структурного синтеза хранилищ гетерогенных данных.

2. Общей проблемой для каждого многоэтапного промышленного предприятия является гетерогенность данных производственных процессов, для анализа которых и решения задач управления качеством продукции необходимо разместить эти данные в едином ХГД.

3. Разработана универсальная методология структурного синтеза единого ХГД, предусматривающая системный анализ технологических процессов с целью разработки информационных сущностей «План производства», «Готовая продукция», «Единица продукции», их онтологического описания, разработки логической структуры единого ХГД и обоснования выбора технологии разработки единого ХГД, которая обеспечит реализацию логической структуры ХГД.

4. На основе методологии структурного синтеза ХГД разработан модуль «ХГД МП», являющийся неотъемлемой частью АС ВИП, и соответствующие программное обеспечение единого ХГД МП, работоспособность которого подтверждена результатами тестовых испытаний.

Рекомендуется дальнейшее развитие темы в направлении применения предложенной в диссертации методологии структурного синтеза ХГД для разработки единых ХГД промышленных предприятий других отраслей и их использование для повышения качества выпускаемой продукции и оптимизации бизнес-процессов предприятия.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИОННОГО ИССЛЕДОВАНИЯ

Статьи, опубликованные в рецензируемых научных журналах и изданиях, определенных ВАК РФ и Аттестационным советом УрФУ:

1. Porshnev S. The Development of a Heterogeneous MP Data Model Based on the Ontological Approach / S. Porshnev, A. Borodin, **O. Ponomareva**, S. Mirvoda and O. Chernova // Symmetry. – 2021. – Vol. 13(5). – No. 813 (1,2 п.л./ 0,24 п.л.) (Scopus, Web of Science)
2. Ponomareva O.A. Date preparation module of automated metallurgical products production system / **O.A. Ponomareva**, S. Porshnev, A. Borodin, & S. Mirvoda // IOP Conference Series: Materials Science and Engineering. – 2021. – Vol. 1047(1). – No. 012003 (0,26 п.л./ 0,06 п.л.) (Scopus)
3. Porshnev S.V. A Mathematical Model for the Description of Metallurgical Production on the Basis of Heterogeneous Data / S. Porshnev, **O. Ponomareva**, S. Trofimov and O. Anchugova // 2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2020. – 2020. – No. 9117711. – pp. 496-498 (0,24 п.л./ 0,06 п.л.) (Scopus)
4. Porshnev S. Problems and Methods for Integrating Heterogeneous Data (on Example, Metallurgical Production). / S. Porshnev, **O. Ponomareva**, A. Borodin, S. Mirvoda // International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon 2019. – 2019. – No. 8934336 (0,49 п.л./ 0,12 п.л.) (Scopus)
5. Borodin A. Design of DSQLM Language Extensions. / A. Borodin, S. Mirvoda, S. Porshnev, **O.A. Ponomareva** // 2019 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2019. – 2019. – No. 8736644. – pp. 295-298 (0,24 п.л./ 0,06 п.л.) (Scopus)

6. Поршнеv С.В. Автоматизированная система выпуска металлургической продукции: от разнородных баз данных к созданию единого хранилища данных / С.В. Поршнеv, **О.А. Пономарева**, А.М. Бородин, С.Г. Мирвода // Автоматизация. Современные технологии. – 2018. – Т. 72(10). – С.435-440 (0,45 п.л. /0,11 п.л.)
7. Поршнеv С.В. Модификация функции penalty R-дерева над обобщенным деревом поиска индексов для повышения производительности модуля CUBE POSTGRESQL / С.В. Поршнеv, **О.А. Пономарева**, А.М. Бородин, С.Г. Мирвода // Программная инженерия. – 2018. – Т. 9(1). – С.22-28 (0,6 п.л. /0,15 п.л.)
8. Porshnev S. The Concept of Automated System of Steel Production Storage Data Structure. / S. Porshnev, A. Borodin, A. Kovaleva, **О.А. Ponomareva** / International Conference on Applied Mathematics and Computational Science, ICAMCS.NET 2018. – 2018. – No. 8955775. – pp. 19-22 (0.35 п.л./ 0,09 п.л.) (Scopus, Web of Science)
9. Borodin A. Improving generalized inverted index lock wait times. / A. Borodin, S. Mirvoda, S. Porshnev, **О.А. Ponomareva** // Journal of Physics: Conference Series. – 2018. – Vol. 944(1). – No. 012022. (0,38 п.л./ 0,09 п.л.) (Scopus, Web of Science)
10. Porshnev S. Computer-Aided system for metallurgy production output: From heterogeneous data bases to unified data warehouse. / S. Porshnev, A. Borodin, E.Dymova, **О.А. Ponomareva** // 2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology, USBEREIT 2018. – 2018. – No. 8384571. – pp. 140-143 (0,24 п.л./ 0,06 п.л.) (Scopus)
11. Trofimov S. Data formalization by algebra of sets in multi-dimension space. / S.Trofimov & **О.А. Ponomareva** // 2016 Dynamics of Systems, Mechanisms and Machines, Dynamics. – 2016. – No. 7819099. (0,37 п.л./ 0,19 п.л.) (Scopus, Web of Science)

В других изданиях:

12. Поршнеv С.В. Обоснование структуры хранения данных АИС ВМП / С.В. Поршнеv, А.М. Бородин, **О.А. Пономарева**, С.Г. Мирвода // Новые информационные технологии в исследовании сложных структур: Материалы 12 конференции. – 2018. – С. 40 (0,69 п.л. /0,17 п.л.)

Перечень сокращений и обозначений

Сокращение	Расшифровка	Сокращение	Расшифровка
ERP-система	EntERPrise Resource Planning	ИС	Информационная система
MES – система	Manufacturing Execution System	КЗ	Конструктор Запросов
АИС МОД	Автоматизированная Информационная Система Моделирования Организационной Деятельности (металлургического предприятия)	МНЛЗ	Машина Непрерывного Литья Заготовок
АИС САД	Автоматизированная Информационная Система Сбора и Анализа Данных (металлургического предприятия)	МП	Металлургическое Производство
АС ВМП	Автоматизированная Система Выпуска Metallургической Продукции	НЛЗ	Непрерывное Литье Заготовок
АСУ ТП	Автоматизированная Система Управления Технологическими Процессами	ОДАСП	(Информационная система) Обмена Данными с Автоматизированными Системами Предприятия
БД	База Данных	ПП	План Производства
ГД	Гетерогенные Данные	ПТП	Параметры технологического процесса
ГП	Готовая Продукция	ФХС	Физико-Химические Свойства
ЕП	Единица Продукции	ХГД МП	Хранилище Гетерогенных Данных Metallургического Производства
ДЕП	триплет «событие, ЕП, источник», описывающий динамику ЕП	ХД	Хранилище Данных
ИИ	Источник информации		