

В. А. Дударев

**ИНТЕГРАЦИЯ
ИНФОРМАЦИОННЫХ
СИСТЕМ в области
НЕОРГАНИЧЕСКОЙ
ХИМИИ
и материаловедения**



URSS

Национальный исследовательский университет
«Высшая школа экономики»
Институт металлургии и материаловедения им. А. А. Байкова
Российской академии наук

В. А. Дударев

**ИНТЕГРАЦИЯ
ИНФОРМАЦИОННЫХ
СИСТЕМ
В ОБЛАСТИ
НЕОРГАНИЧЕСКОЙ
ХИМИИ
И МАТЕРИАЛОВЕДЕНИЯ**



URSS
МОСКВА

Дударев Виктор Анатольевич

Интеграция информационных систем в области неорганической химии и материаловедения. — М.: КРАСАНД, 2017. — 320 с.

Книга посвящена вопросам создания интегрированных информационных систем в области неорганической химии. В работе приведен краткий обзор наиболее значимых фактографических баз данных по свойствам неорганических веществ, созданных в мире, и выполнен анализ архитектуры современных информационных систем в области материаловедения. Рассмотрены основные методы создания интегрированных систем (ETL, ЕП, ЕАИ), и на их основе предложена комплексная методология интеграции материаловедческой информации, учитывающая требования пользователей и программ анализа данных. Приведены особенности реализации интегрированной информационной системы по свойствам неорганических веществ Института металлургии и материаловедения им. А. А. Байкова Российской академии наук (ИМЕТ РАН). В заключении рассмотрены примеры использования консолидированной информации для поиска закономерностей в данных и компьютерного конструирования новых перспективных неорганических соединений.

Книга предназначена для специалистов, занимающихся разработкой информационных систем по свойствам неорганических веществ. Изложенный материал позволяет грамотно реализовать структуры хранения информации, что открывает пути для дальнейшей интеграции с другими информационными системами и проведения интеллектуального анализа накопленных данных с целью прогнозирования свойств неорганических веществ.

Издательство «КРАСАНД». 117335, Москва, Нахимовский пр-т, 56.
Формат 60×90/16. Печ. л. 20. Доп. тираж.

Отпечатано в ООО «Курганский Дом печати». 640022, Курган, ул. К. Маркса, 106.

ISBN 978–5–396–00745–1

© КРАСАНД, 2016

20947 ID 220601



9 785396 007451



Все права защищены. Никакая часть настоящей книги не может быть воспроизведена или передана в какой бы то ни было форме и какими бы то ни было средствами, будь то электронные или механические, включая фотокопирование и запись на магнитный носитель, а также размещение в Интернете, если на то нет письменного разрешения владельца.

Оглавление

ВВЕДЕНИЕ	8
ГЛАВА 1. ОСОБЕННОСТИ ПРИНЯТИЯ РЕШЕНИЙ ПРИ ПРОГНОЗИРОВАНИИ СВОЙСТВ НЕОРГАНИЧЕСКИХ ВЕЩЕСТВ	10
1.1. Способы конструирования неорганических соединений	12
1.1.1. Методы квантовой механики	13
1.1.2. Простейшие эмпирические зависимости.....	14
1.1.3. Многомерные классифицирующие правила.....	15
1.2. Математические методы распознавания	18
1.2.1. Формальная постановка задачи прогнозирования	19
1.2.2. Методы обучения ЭВМ распознаванию образов	22
1.2.3. Способы повышения достоверности прогнозов	32
Краткие выводы	40
ГЛАВА 2. АНАЛИЗ АРХИТЕКТУРНЫХ ОСОБЕННОСТЕЙ ИНФОРМАЦИОННЫХ СИСТЕМ ПО СВОЙСТВАМ НЕОРГАНИЧЕСКИХ ВЕЩЕСТВ	41
2.1. Обзор ИС СНВМ для электроники.....	41
2.2. Создание ИС по информационным ресурсам неорганической химии «IRIC».....	71
2.2.1. Схема данных	72
2.2.2. Web-приложение	74
2.3. Архитектура современных информационных систем по свойствам веществ	78
2.3.1. Использование трехзвенной архитектуры	78
2.3.2. Недостатки ИС СНВМ	81
2.3.3. Обобщенная структура данных для ИС СНВМ	82

2.4. Информационные системы по свойствам неорганических веществ ИМЕТ РАН.....	85
2.4.1. Разработка ИС по веществам с особыми акустооптическими, электрооптическими и нелинейнооптическими свойствами «Кристалл»	85
2.4.2. Разработка ИС по ширине запрещенной зоны неорганических соединений «Bandgap»	94
2.4.3. ИС по свойствам неорганических соединений «Фазы»	97
2.4.4. ИС по фазовым диаграммам систем с полупроводниковыми фазами «Диаграмма».....	98
2.4.5. ИС по свойствам кремния «Кремний»	99
2.4.6. Разработка программного комплекса для удаленного администрирования гетерогенных БД ИМЕТ РАН.....	99
2.4.7. Особенности ИС ИМЕТ РАН.....	104
2.5. Расчетные подсистемы информационных систем по свойствам неорганических веществ	104
Краткие выводы	109

Глава 3. СИСТЕМНЫЙ ПОДХОД К ИНТЕГРАЦИИ ИНФОРМАЦИОННЫХ СИСТЕМ 110

3.1. Методы интеграции гетерогенных информационных систем.....	110
3.1.1. Актуальность интеграции	110
3.1.2. Проблемы при интеграции информационных систем	111
3.1.3. Методы интеграции ИС.....	112
3.1.4. Проблемы при интеграции гетерогенных источников информации.....	125
3.2. Системный анализ методов интеграции	126
3.2.1. Базовые информационные процессы в локальных ИС.....	126
3.2.2. Метод интеграции корпоративной информации ЕП.....	128
3.2.3. Метод интеграции на основе хранилищ данных ETL	130

3.2.4. Интеграция корпоративных приложений EAI.....	133
3.2.5. Обобщенная схема методов интеграции гетерогенных информационных систем	135
3.3. Методология интеграции информационных систем....	140
3.4. Интеграция гетерогенных источников данных информационных систем.....	147
3.4.1. Разрешение платформенных и системных конфликтов.....	148
3.4.2. Разрешение синтаксических и структурных конфликтов	151
3.4.3. Разрешение семантических конфликтов	155
3.5. Платформа для разработки интегрированной ИС СНВМ.....	161
3.5.1. Производительность	162
3.5.2. Безопасность	165
3.5.3. Надежность	167
3.5.4. Интероперабельность	168
3.5.5. Совокупная стоимость владения	170
Краткие выводы	172
ГЛАВА 4. СИСТЕМНЫЙ ПОДХОД К РАЗРАБОТКЕ ХРАНИЛИЩА ДАННЫХ ПО СВОЙСТВАМ НЕОРГАНИЧЕСКИХ ВЕЩЕСТВ ДЛЯ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ	173
4.1. Диаграммы потоков данных DFD.....	173
4.2. Методология функционального моделирования IDEF0.....	174
4.3. ER-модель хранилища данных.....	176
4.4. Реляционная структура ХД	177
4.5. Извлечение, преобразование и загрузка данных в ХД	179
4.5.1. Процедура извлечения.....	179
4.5.2. Процедура преобразования данных	182
4.5.3. Процедура загрузки	186
Краткие выводы	187

ГЛАВА 5. ИСПОЛЬЗОВАНИЕ ВИРТУАЛЬНОЙ ИНТЕГРАЦИИ ДАННЫХ ПРИ ПРОГНОЗИРОВАНИИ СВОЙСТВ НЕОРГАНИЧЕСКИХ ВЕЩЕСТВ	188
5.1. Подходы к интеграции информации средствами ЕИ	188
5.2. Реализация интеграции гетерогенных источников данных информационных систем	193
5.2.1. Описание структуры метабазы	193
5.2.2. Расчет достоверности информации, основанный на экспертных оценках	201
5.2.3. Разработка программных адаптеров интегрируемых информационных систем	202
5.2.4. Разработка предметного посредника	207
Краткие выводы	212
ГЛАВА 6. ИСПОЛЬЗОВАНИЕ ИНТЕГРАЦИИ ПРИЛОЖЕНИЙ ДЛЯ ИНФОРМАЦИОННОЙ ПОДДЕРЖКИ СПЕЦИАЛИСТОВ В ОБЛАСТИ НЕОРГАНИЧЕСКОЙ ХИМИИ	213
6.1. Интеграция распределенных гетерогенных Web-приложений информационных систем	213
6.2. Реализация интеграции гетерогенных Web-приложений информационных систем	222
6.2.1. Описание структуры метабазы	222
6.2.2. Загрузка информации в метабазу	229
6.2.3. Поиск релевантной информации по содержимому метабазы	231
6.2.4. Осуществление безопасного перехода пользователя между Web-приложениями интегрируемых информационных систем	233
6.3. Единая точка входа в ИС СНВМ	237
6.3.1. Поиск релевантной информации	238
6.3.2. Разработка Web-приложения ИС	239
6.4. Создание системы единой авторизации	244
Краткие выводы	251

ГЛАВА 7. ПРИМЕНЕНИЕ ИНТЕГРИРОВАННОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ПОИСКА ЗАКОНОМЕРНОСТЕЙ И КОМПЬЮТЕРНОГО КОНСТРУИРОВАНИЯ НОВЫХ СОЕДИНЕНИЙ	252
7.1. Интерполяция неизвестных значений в обучающих выборках	253
7.1.1. Краткий обзор методов заполнения пропусков в данных	253
7.1.2. Методика заполнения неизвестных значений с учетом специфики предметной области	255
7.2. Этапы компьютерного конструирования новых соединений	262
7.3. Перспективные полупроводники ABX_2	265
7.4. Перспективные диэлектрики $A_2B_2(XO_4)_3$	269
7.5. Прогноз образования сегнетоэлектрических хлоридов A_2BCl_4	271
7.6. Прогноз образования соединений состава AB_2X_4	275
Краткие выводы	290
ЗАКЛЮЧЕНИЕ	292
ЛИТЕРАТУРА	293
ПРИЛОЖЕНИЕ	312