

В.И. Аверченков, А.В. Заболеева-Зотова, Ю.М. Казаков,
Е.А. Леонов, С.М. Рощин

СИСТЕМА ФОРМИРОВАНИЯ ЗНАНИЙ В СРЕДЕ ИНТЕРНЕТ

Монография

4-е издание, стереотипное

Москва
Издательство «ФЛИНТА»
2021

УДК 004.415.2.045:004.738.5
ББК 73я73
А19

Аверченков В.И.

А19 Система формирования знаний в среде Интернет: монография / В.И. Аверченков, А.В. Заболеева-Зотова, Ю.М. Казаков, Е.А. Леонов, С.М. Рощин. – 4-е изд., стер. – Москва : ФЛИНТА, 2016. – 181 с. – ISBN 978-5-9765-1266-5. – Текст : электронный

Проанализированы проблемы разработки и применения систем и модулей поиска информации, основанных на принципах искусственного интеллекта, а также формирование предметно-ориентированных порталов, объединяющих систематизированные информационные ресурсы сети Интернет конкретной предметной области, механизмы их обработки. Разработаны математические модели и алгоритм для систем мониторинга информации на основе онтологического подхода, описана концепция системы формирования знаний.

УДК 004.415.2.045:004.738.5
ББК 73я73

ISBN 978-5-9765-1266-5

© Коллектив авторов, 2016
© Издательство «ФЛИНТА», 2016

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
1. АНАЛИЗ ПРОБЛЕМ РАЗРАБОТКИ И ПРИМЕНЕНИЯ МЕТОДОВ СИСТЕМНОГО АНАЛИЗА ОБРАБОТКИ РАСПРЕДЕЛЕННОЙ ПРОБЛЕМНО-ОРИЕНТИРОВАННОЙ ИНФОРМАЦИИ В СРЕДЕ ИНТЕРНЕТ	5
1.1. Роль мониторинга и системного анализа информации в среде Интернет	5
1.2. Проблемы поиска теоретической научной информации в Интернет	6
1.3. Формализованное описание и классификация методов мониторинга и системного анализа информации	11
1.4. Обзор существующих систем автоматизации мониторинга и системного анализа распределенной информации	18
1.5. Анализ подходов к организации хранилищ данных	21
2. КОНЦЕПЦИЯ МОНИТОРИНГА И СИСТЕМНОГО АНАЛИЗА РАСПРЕДЕЛЕННОЙ ПРОБЛЕМНО-ОРИЕНТИРОВАННОЙ ИНФОРМАЦИИ В СРЕДЕ ИНТЕРНЕТ	24
2.1. Система формирования знаний	24
2.2. Структура и обобщенный алгоритм работы системы формирования знаний	26
2.3. Представление знаний о предметной области в системе формирования знаний	30
2.4. Формирование онтологии предметной области	33
2.5. Разработка структуры многоагентной системы доступа к информационным ресурсам научного и образовательного назначения	44
3. СТРУКТУРА И ПРИНЦИПЫ РАБОТЫ МОДУЛЯ МОНИТОРИНГА WEB-РЕСУРСОВ	54
3.1. Объектная модель сервиса сканирования	55
3.2. Описание структуры конфигурационных и вспомогательных файлов	57
3.3. Обобщенный алгоритм работы модуля сканирования	62
3.4. Информационное обеспечение системы управления модуля мониторинга	68
3.5. Алгоритм создания запросов к поисковым системам	69
3.6. Алгоритм разбора страниц, получаемых от поисковых систем	71
3.7. Интерфейс системы модуля мониторинга	73
3.8. Администрирование и управление информационными ресурсами для хранилища данных web-ресурсов	82
4. СТРУКТУРА И ПРИНЦИПЫ РАБОТЫ МОДУЛЕЙ АНАЛИТИКИ, ИНДЕКСИРОВАНИЯ И ПОИСКА	85
4.1. Системный анализ распределенной проблемно-ориентированной информации	85
4.2. Семантический анализ текста	86
4.3. Теоретико-методологические основы лексической семантики	87
4.4. Текст как объект лингвистического исследования	93
4.4.1. Структура текста	94
4.4.2. Методы лингвистического анализа текста	96
4.4.3. Представление структуры текста в виде дискурсного графа	101
4.5. Средства компьютерного представления Естественно-Языковых феноменов	102
4.5.1. Лингвистическая обработка ЕЯ	103
4.5.2. Морфологический уровень	104
4.5.3. Синтаксический уровень	107
4.5.4. Семантический уровень	109
4.5.5. Основные лингвистические теории	110
4.5.6. Проблема многозначности	114
4.6. Обобщенная архитектура системы семантического анализа текста	116

4.6.1.	Классификация ЕЯ-систем.....	116
4.6.2.	Функциональная структура системы	116
4.6.3.	Модели представления знаний в ЕЯ-системах	117
4.6.4.	Представление семантической информации.....	122
4.7.	Подсистемы индексирования, кластеризации информации.....	128
5.	РАЗРАБОТКА ПРОБЛЕМНО-ОРИЕНТИРОВАННОГО WEB-ПОРТАЛА	150
5.1.	Организация портала знаний на основе онтологий.....	150
5.2.	Структура сервисов проблемно-ориентированного web-портала	155
	Заключение	170
	СПИСОК ЛИТЕРАТУРЫ.....	171