

Информационно-  
аналитическая  
система



**Cognitive**  
Technologies Ltd

AST@RTA

# Предпосылки создания ИАС АСТАРТА

- ❑ Высокая потребность в оперативной информации для принятия управленческих решений
- ❑ Необходимость проведения исследований рынков любой отрасли и тематики
- ❑ Постоянно увеличивающийся поток разнородной информации
- ❑ Проблемы с систематизацией, хранением и поиском информации как в бумажном, так и в электронном виде
- ❑ Большое количество рутинных операций
- ❑ Как следствие - снижение оперативности работы экспертов
- ❑ Высокая степень субъективности экспертных оценок
- ❑ Отсутствие преемственности экспертной деятельности

# Решаемые задачи

- ❑ **Оперативный мониторинг бумажных и электронных СМИ**
- ❑ **Обеспечение автоматической систематизации и рубрикации поступающей информации**
- ❑ **Исследование предметной области любого научно-технического направления**
- ❑ **Анализ тенденций и прогнозирование развития рынка продукции и услуг**
- ❑ **Анализ структуры рынка и конкурентная разведка**
- ❑ **Предоставление экспертам эффективного инструментария для анализа накопленной информации и разработки рекомендаций для руководства**
- ❑ **Создание полнотекстового архива документов, с возможностью их быстрого поиска по тексту и реквизитам**
- ❑ **Обеспечение простоты доступа экспертов ко всем хранимым информационным материалам**
- ❑ **Обеспечение преемственности экспертной деятельности путем максимально возможной формализации экспертных знаний**

# Функциональные особенности

- ❑ Построение пользовательского иерархического рубрикатора по любой предметной области
- ❑ Автоматизированное обучение и переобучения системы в процессе работы
- ❑ Автоматическое построение списка терминов и алгоритмов, определяющих принадлежность к рубрике и возможность их ручной корректировки
- ❑ Ввод бумажных документов с использованием технологии распознавания текстов CuneiForm 2000
- ❑ Автоматическое слежение за публикацией «свежих» новостей на Интернет-СМИ с указанием расписания их опроса
- ❑ Встроенный редактор для настройки на новые новостные сайты
- ❑ Автоматический разбор формата новостной ленты
- ❑ Автоматическое слежение за документами в директориях, указанных пользователем
- ❑ Встроенные средства поиска и просмотра информации в архиве
- ❑ Встроенный генератор отчетов и дайджестов
- ❑ Средства статистического анализа поступающей информации
- ❑ Встроенная СУБД

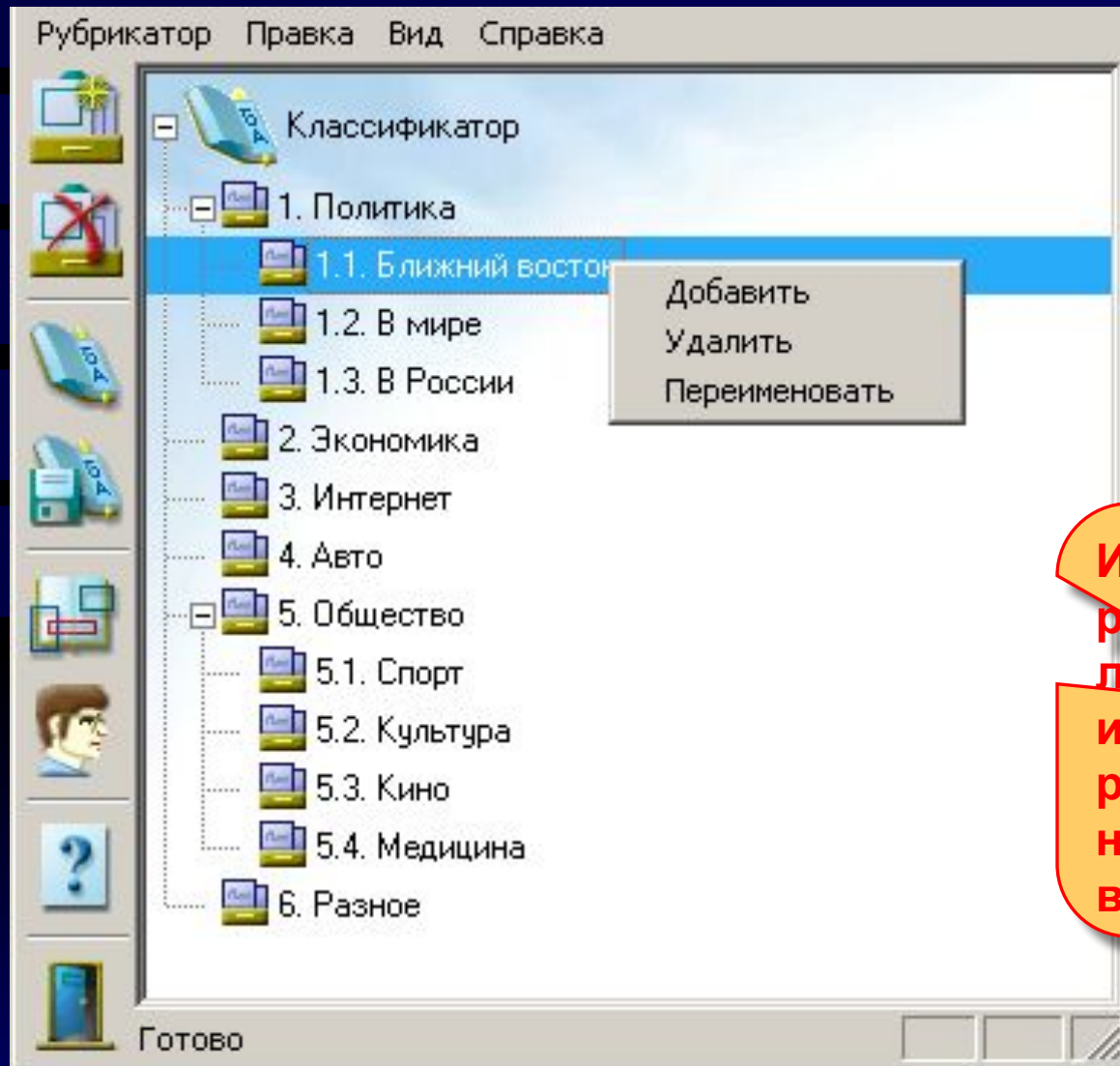
# Функциональная схема ИАС АСТАРТА



# Технологическая схема работы ИАС АСТАРТА

- Создание тематического рубрикатора
- Обучение авторубрикатора
- Ввод информации с автоматическим определением тематики документов
- Поиск
- Формирование дайджестов
- Статистические отчеты

# Тематический рубрикатор



**ИАС АСТАРТА**  
работает как с  
линейным так и с  
иерархическим  
рубрикатором  
неограниченной  
вложенности

# Обучение рубрикатора

№ (25)	Имя документа
16	rian_0187
17	rian_0188
18	rian_0200
19	rian_0201
20	rian_0203
21	rian_0212
22	rian_0219
23	rian_0220
24	rian_0230
25	rian_0248

ТЕГЕРАН, 14 апреля. /Корр. РИА  
Николай Терехов/. Амери  
арестовали в Ираке корр  
телеканала "Аль-Алем" А  
Как сообщило в понедель  
ИРНА, журналист занима  
материалов о ситуации в с  
снял на камеру американские танки, входящие в  
иракский город Эль-Кут.  
Военнослужащие США забрали у журналиста  
кассету со снятыми материалами, арестовали и  
увезли в неизвестное место, сообщило ИРНА.  
Иракский город Эль-Кут находится в 120

**Обучение каждой рубрики осуществляется автоматически с помощью обучающих выборок**



# Обучение рубрикатора

Эксперт. Рубрикатор 'Классификатор'

Классификатор

- 1. Политика
  - 1.1. Ближний восток
  - 1.2. В мире
  - 1.3. В России
- 2. Экономика
- 3. Интернет
- 4. Авто
- 5. Общество
- 6. Разное

Термин (2)	Оценка
Сирия	3
ирак	5

Термин (58)	
ирак	
иракский город	
иракский столица	
каир	
кандидатура	
катарский телеканал	
катарский	
корреспондент телеканал	186
население	115
окрчжение	117

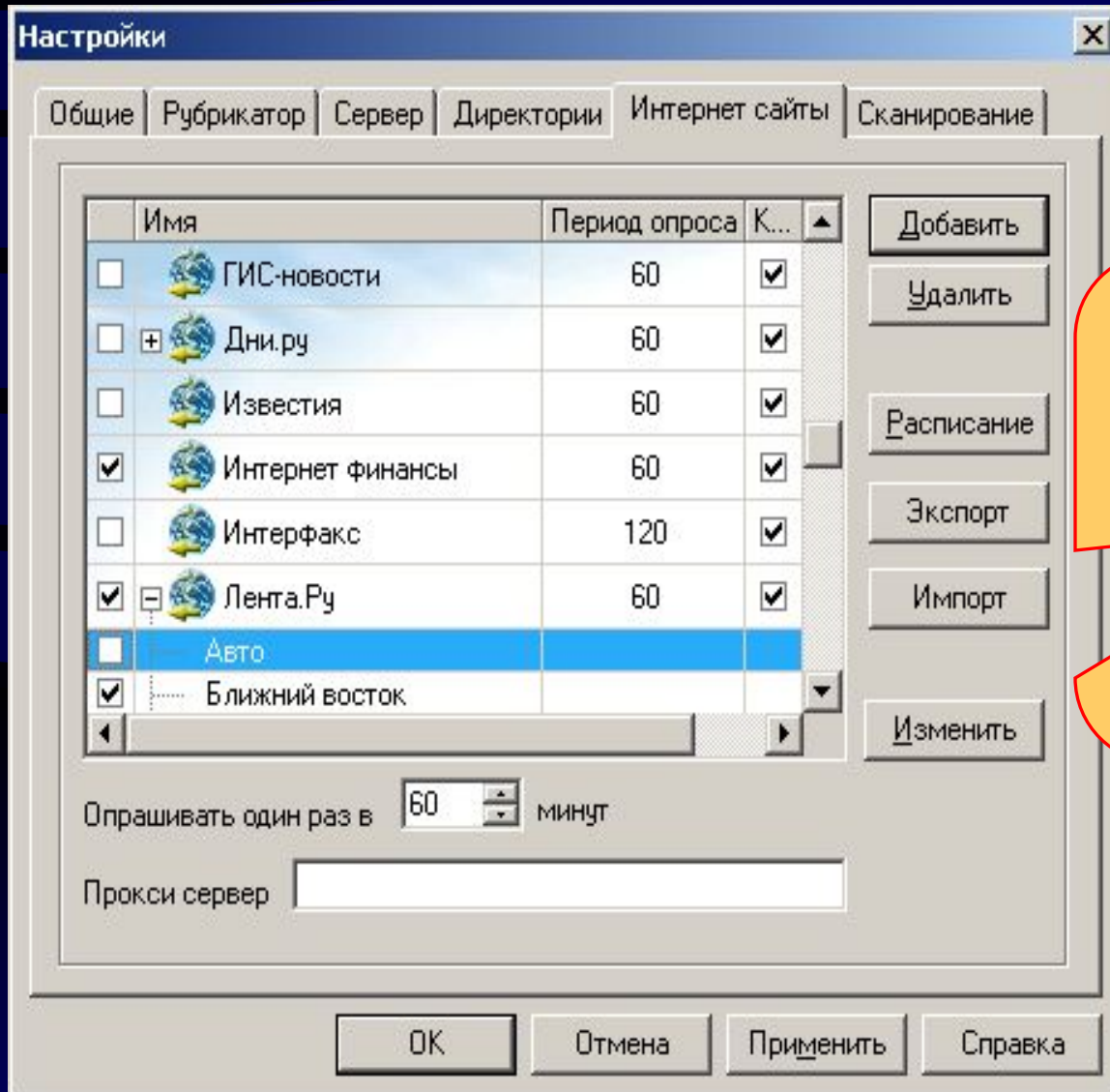
**В дальнейшем эксперт может осуществить ручную корректировку автоматических настроек рубрикатора**

# Ввод информации

**ИАПК обеспечивает автоматический сбор информации из следующих источников:**

- Интернет-СМИ
- ленты новостных агентств
- специализированные патентные сайты
- бумажные документы

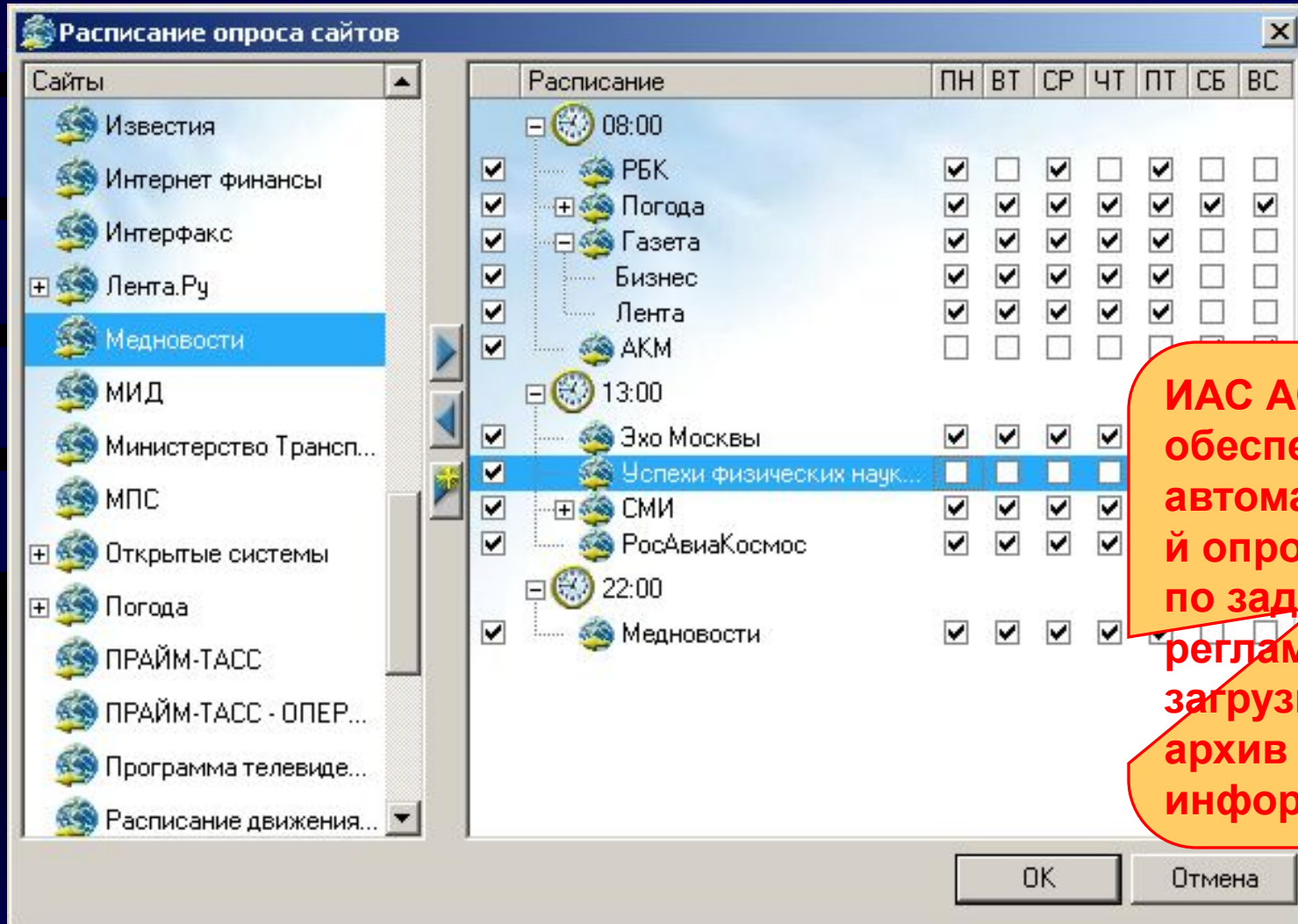
# Работа с Интернет-СМИ



**Менеджмент сайтов:**

- добавление
- удаление
- корректировка
- экспорт/импорт описания сайта

# Работа с Интернет-СМИ



# Работа с Интернет-СМИ

Сайт "АКМ"

Название: АКМ

Период опроса: 60 минут Последнее обновление: 2003-04-14 19:33

Таймаут: 30 секунд Количество повторов: 10

Имя папки:

Маска файлов: akm

Кодировка: Автоматическая Специальные настройки...

Ссылка = `http://www.akm.ru/rus/info/todaynews.stm`

РегВыр ограничения индексной страницы = ""

РегВыр индексной страницы = `"<tr valign=\"\"top\">.<td><b>([0-9]+)`

Индексные поля = HOUR, MINUTE, URL, Тема

РегВыр новостной страницы = `"<td valign=\"\"top\"><b>([0-9]+)`

Поля новости = DAY, MONTH, YEAR, HOUR, MINUTE, TEXT

Имя Документа = Поле(Тема)

OK Отмена

Встроенный редактор позволяет описать и подключить дополнительные новостные сайты

# Ввод бумажных документов

- ❑ Использование новейших уникальных технологий, разработанных российскими учеными: самообучающийся Omnifont, адаптивное распознавание, нейронные сети, когнитивный анализ™, меридианальная сегментация™
- ❑ Ввод печатного и машинописного текста, факсов, графики (черно-белой и цветной), таблиц любой сложности
- ❑ Распознавание и словарный контроль 23 основных языков
- ❑ Поддержка всех TWAIN-совместимых сканеров
- ❑ Использование ADF, сетевое сканирование, потоковое распознавание
- ❑ Сохранение сложного форматирования документа
- ❑ Встроенный редактор
- ❑ Сохранение результатов в распространенных форматах (\*.rtf, \*.dbf, \*.html)



# Рабочий стол системы

Файл Вид Сервис Справка

- Аккумулятор...
- Газета
- Газета - Авто
- ГИС-новости
- Известия
- Интернет фи...
- Интерфакс
- Лента.Ру**
- Медновости
- МИД
- Открытые с...
- ПРАЙМ-ТАСС
- ПРАЙМ-TAC...
- РБК
- Регионы.ру
- РИА 'Новости'
- РИА 'Новост...
- СМИ
- Эхо Москвы

№ (73)	Тема	Дата-Время п...
8	Кувейт собирает страны Залива для об...	02:41 07.04.2003
9	Великобритания выдворяет р...	
10	Подозреваемый "вашингтонс...	
11	Саддам заплатит 15 миллионов за кажд...	03:50 07.04.2003
12	Британцы защитятся от огня союзнико...	04:12 07.04.2003
13	При штурме Басры погибли трое британ...	04:42 07.04.2003
14	Британский водитель напал с ножом на...	05:19 07.04.2003
15	Гран-при Бразилии выиграл Кими Райк...	05:31 07.04.2003

**Состав выбранной папки с документами**

**Содержание выделенного документа**

Реквизиты документа "Гр..."

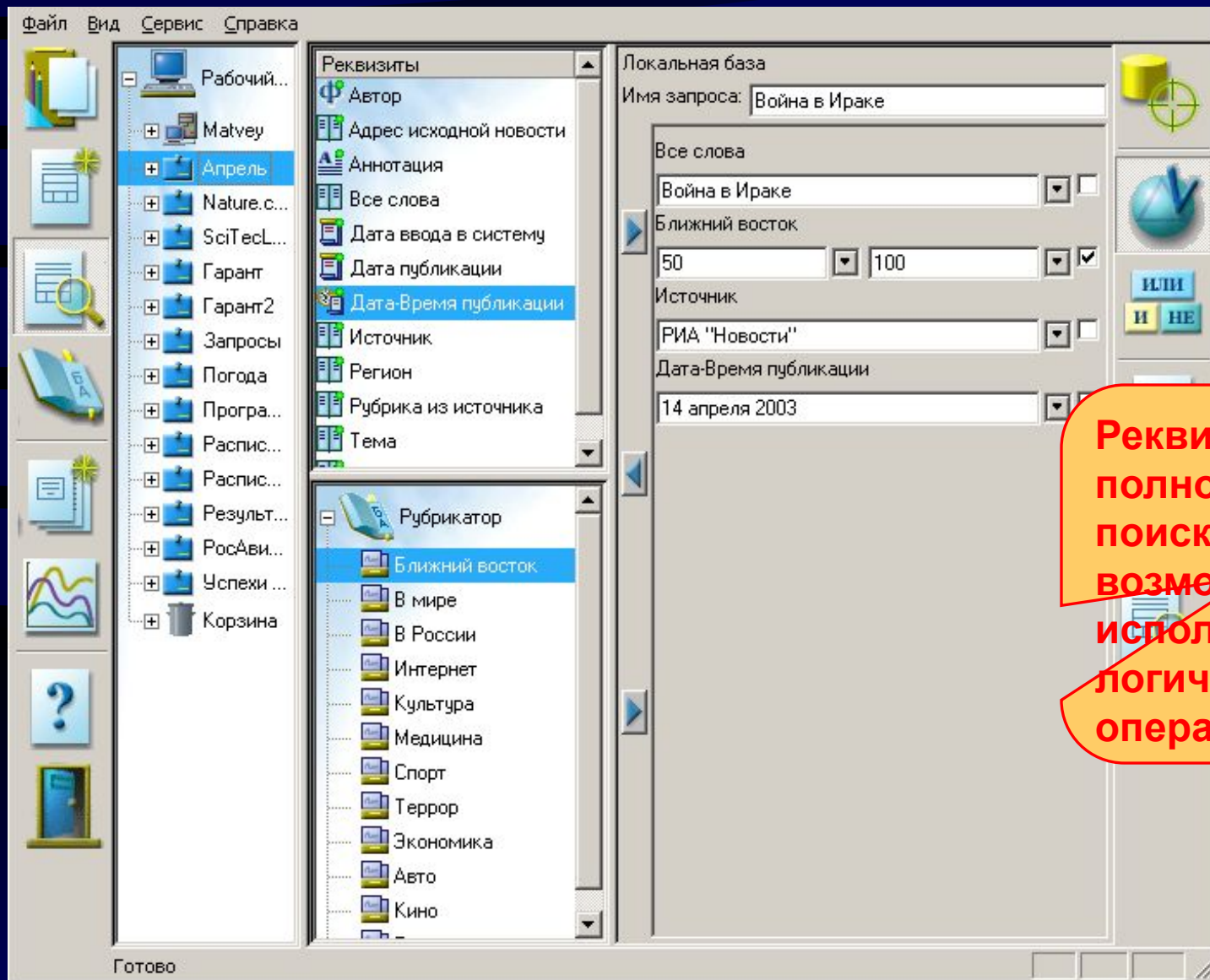
Название	Значение
Дата ввода в ...	07.04.2003
Дата публика...	07.04.2003
Дата-Время п...	05:31 07.04.2003
Источник	www.lenta.ru
Рубрика из ис...	Спорт
Тема	Гран-при Бразилии
Формат	Лента.Ру
Спорт	55

Материалы могут быть представлены по датам или по рубрикам

Реквизиты выделенного документа

...ий этап чемпионата мира  
...нский пилот команды  
...ен. Второе место занял  
"Джордана", а третьим  
"Рено", который не смог  
...рчине аварии. Гонка,

# Поиск документов



**Реквизитный и полнотекстовый поиск с возможностью использования логических операций**



# Формирование дайджестов

- ❑ ИАС АСТАРТА обеспечивает автоматическое формирование дайджестов из материалов, хранящихся в системе
- ❑ Могут быть сформированы дайджесты:
  - ❑ полнотекстовые или только по аннотации
  - ❑ многотомные
  - ❑ включающие только оглавление
  - ❑ отсортированные с учетом разбивки документов по рубрикам и по значениям реквизитов
- ❑ Формирование дайджеста осуществляется с учетом пользовательских настроек шаблона, что обеспечивает единый стиль документа

# Формирование дайджестов

Создание дайджеста по рубрике "В мире"

Формат: MS Word (\*.doc)

Не печатать картинки

Печатать только оглавление

Пронумеровать оглавление

Печатать оглавление заглавными буквами

Разбить на отдельные тома при превышении 500 стр.

Сохранить оригинальное форматирование текста

Сортировать по рубрикам

Отчет по аннотации Аннотация

Порядок сортировки по реквизитам

- Тема
- Автор
- Дата-Время
- Аннотация
- Дата ввода
- Дата публикации
- Источник
- Регион
- Рубрика из источника
- Формат
- Адрес исходной новости

Шаблоны

- Отчет по аннотации
- Программа телевидения

Пользовательские шаблоны дайджеста

Верхний колонтитул

Нижний колонтитул

Оглавление

<Тема>

Текст

Верхний колонтитул

Страница №<Номер страницы> из <Всего страниц>

Заголовок

<Тема>

Обычный текст

Таблица

Подпись

<Формат>

<Источник>

Нижний колонтитул

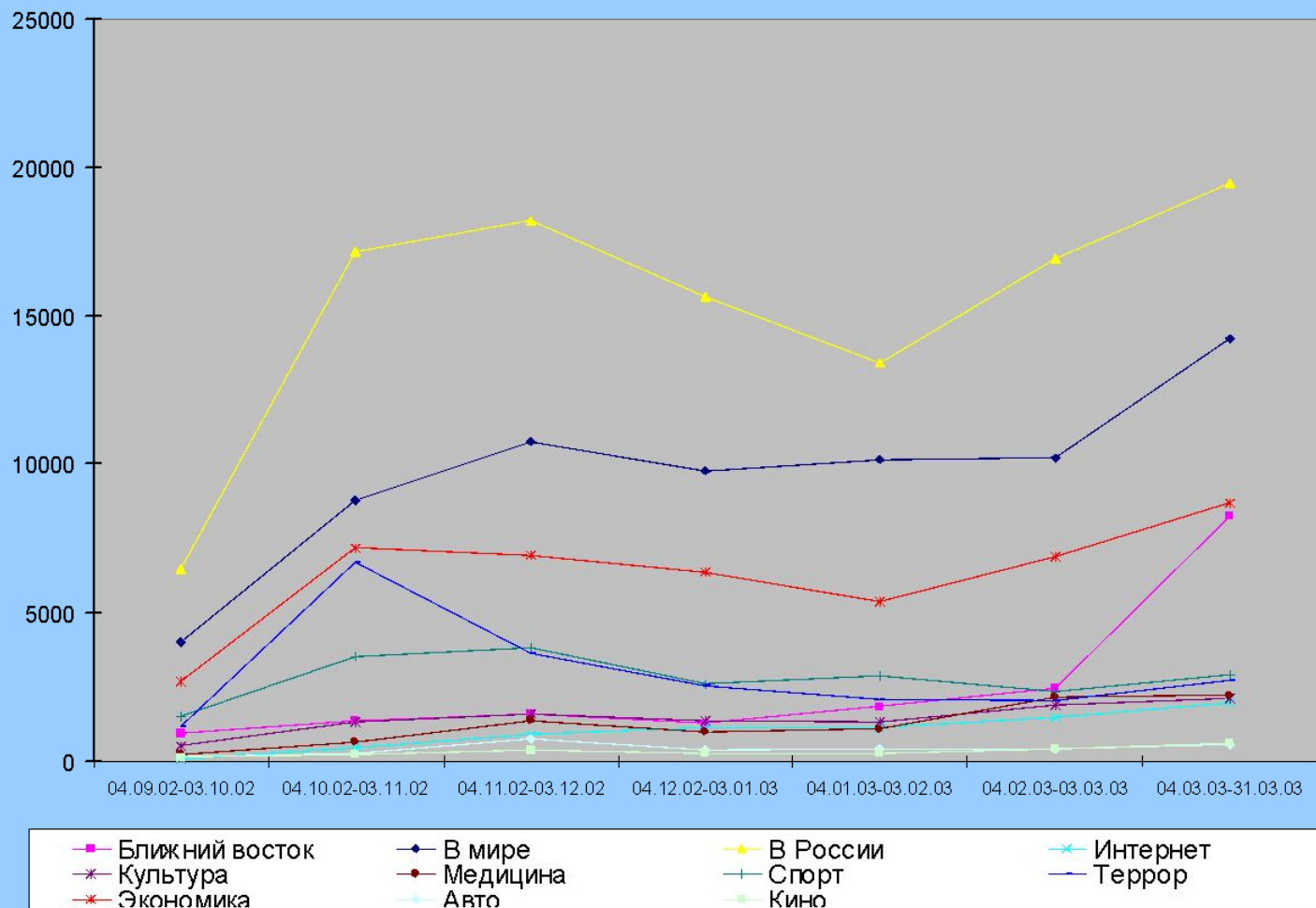
Рубрики

Настройка порядка сортировки по реквизитам

Описание шаблона

# Статистический анализ

Для качественной оценки исследуемой величины используется пакет MS Excel



# Результаты внедрения

- ❑ **Обеспечение стабильности информационных потоков и возможность расширения числа контролируемых источников информации**
- ❑ **Повышение эффективности использования людских ресурсов и повышение отдачи работы сотрудников**
- ❑ **Возможность и удобство проведения аналитических исследований материалов СМИ высокого уровня сложности**
- ❑ **Систематизация хранения накопленной информации**
- ❑ **Возможность быстрого поиска, сложные запросы в горизонтальном и/или в вертикальном разрезах**
- ❑ **Обеспечение преемственности экспертной деятельности**

# Дополнительная информация

## Cognitive Technologies Ltd.

### Офис компании

*Москва, проспект 60-летия Октября, 9  
тел. 135-55-10, 135-89-68, 135-50-88*



# [www.cognitive.ru](http://www.cognitive.ru)