

Санкт-Петербургский Государственный Университет
Математико-механический факультет

Кафедра информатики

Зернов Алексей Викторович

Разработка системы автоматического
анализа новостных публикаций на
финансовом рынке

Бакалаврская работа

Научный руководитель:
к. ф.-м. н., доцент Григорьев Д. А.

Рецензент:
позиция рецензента рецензент

Санкт-Петербург
2017

SAINT-PETERSBURG STATE UNIVERSITY
Faculty of Mathematics and Mechanics

Computer Science Department

Alexey Zernov

Development of the automatic analysis system of a financial market's news publications

Bachelor's Thesis

Scientific supervisor:
assistant professor Dmitry Grigoryev

Reviewer:
reviewer position reviewer

Saint-Petersburg
2017

Оглавление

Введение	4
1. Обзор существующих решений	5
1.1. IBM: Watson Developer Cloud	5
1.1.1. News Intelligence	5
1.1.2. Social Customer Care	5
1.1.3. News Explorer	5
1.1.4. Investment Advisor	5
1.2. Медиалогия	5

Введение

1. Обзор существующих решений

1.1. IBM: Watson Developer Cloud

1.1.1. News Intelligence

Одним из наиболее интересных сервисов для анализа новостных публикаций является приложение News Intelligence¹.

Приложение предлагает ввести название интересующей Вас компании, после чего предоставляются следующие результаты:

- Наиболее упоминаемые сущности (люди, темы, компании)
- Наиболее просматриваемые новостные публикации
- Анализ тональности новостных публикаций из десяти случайных источников
- Совместные упоминания и оценка их тональности

Важно отметить, что приложение News Intelligence является лишь примером использования данного инструмента, а не готовым продуктом.

1.1.2. Social Customer Care

Еще одним

1.1.3. News Explorer

1.1.4. Investment Advisor

1.2. Медиалогия

Медиалогия² — разработчик автоматической системы мониторинга и анализа СМИ в режиме реального времени.

¹<https://discovery-news-demo.mybluemix.net/>

²<http://www.mlg.ru/>

Данная платформа предоставляет такие решения, как мониторинг СМИ компании (ее брендов, конкурентов и др.) и анализ СМИ и сообщений с использованием уникальной технологии лингвистического анализа текстов.

Из предоставленных примеров отчетов³ на сайте можно увидеть, что сервис учитывает следующее:

- **Количество упоминаний.** Отслеживается динамика по кварталам и месяцам. Отслеживая динамику по дням, платформа определяет наиболее заметные информационные поводы, вызывающие более сильный всплеск упоминаний. Также учитывается цитируемость, совместные упоминания и распределение по тематическим рубрикам.
- **Качество упоминаний.** В рассмотренном отчете предоставлена информация о положительных и негативных сообщениях. Выделены пики и проанализирована связь со СМИ, которые способствуют больше благоприятному или отрицательному всплеску упоминаний.

Также в отчете были учтены распределения по уровням СМИ, по их географическому расположению, и прочее. Однако среди всего этого наиболее важным моментом является как раз анализ текста новости: является упоминание положительным или отрицательным, как это совместно упоминается с другими запросами и тому подобное.

³В качестве образца был взят аналитический отчет компании «Вымпелком» в СМИ за II квартал 2009 года