УДК 004.912

ИНФОРМАЦИОННАЯ МОДЕЛЬ УПРАВЛЕНИЯ МАРКЕТИНГОВОЙ ДЕЯТЕЛЬНОСТЬЮ ОБРАЗОВАТЕЛЬНОЙ ОРГАНИЗАЦИИ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЙ SMM И АВТОМАТИЗИРОВАННОГО ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА СОЦИАЛЬНЫХ СЕТЕЙ

Горбушин Д.А., Гринченков Д.В., Нгуен Фук Хау

Южно-Российский государственный политехнический университет (НПИ) имени М.И. Платова, Новочеркасск, e-mail: gorinwww@gmail.com, grindv@yandex.ru, phuchauptit@gmail.com

В статье рассмотрен вопрос интернет-маркетинга в области профессионального образования. Для решения представленной задачи предлагается применение технологий Social media marketing. Описана стратегия Social media marketing в сфере образовательных организаций высшего профессионального образования. В рамках данной стратегии особое внимание уделяется этапу аналитики, позволяющему получить ценную информацию, которая может быть руководством при принятии решений в процессе управления деятельностью образовательной организации. В качестве анализируемого материала предлагается использовать сообщения пользователей социальных сетей. В качестве методики анализа предлагается использование математических методов интеллектуального анализа тональности текста. Разработана информационная модель управления деятельностью университета с использованием технологий Social media marketing, анализа социальных сетей и анализа тональности мнений. Описанная модель наглядно показывает взаимосвязи целевой аудитории, рынка образовательных услуг и рынка труда. Описана проблематика существующих технологий извлечения информации из социальных сетей и методик интеллектуального анализа тональности текста.

Ключевые слова: анализ тональности текста, компьютерная лингвистика, анализ социальных сетей, интеллектуальный анализ данных, интернет-маркетинг, social media marketing, SMM, высшее профессиональное образование, информационная модель, рынок образовательных услуг, рынок труда

INFORMATION MODEL OF MANAGEMENT OF THE MARKETING ACTIVITY IN EDUCATIONAL INSTITUTION USING SMM TECHNOLOGIES AND INTELLIGENT ANALYSIS OF SOCIAL NETWORKS

Gorbushin D.A., Grinchenkov D.V., Nguyen Phuc Hau

Platov South-Russia State Polytechnic University (NPI), Novocherkassk, e-mail: gorinwww@gmail.com, grindv@yandex.ru, phuchauptit@gmail.com

The issue of internet marketing in the institutions of higher education is studied in the article. Social media marketing technologies are offered to solve the task, paying special attention to the analysis stage. It is offered to consider social networks users' messages as analytical material. Intellectual sentiment analysis is offered as a method for such analyses. Approaches to the solution of a sentiment analysis problem are considered. The perspective application methods of collecting and the intellectual analysis of information necessary for the feedback organization and ensuring high-quality educational services are described. The information model of management of the University activity using Social media marketing techniques, social network analysis and analysis are developed. The described model clearly shows the interrelations of the target audience, the market of educational services and the labor market. The problems of existing technologies for extracting information from social networks and techniques for the intellectual analysis of the tonality of the text are described.

Keywords: sentiment analysis, computational linguistic, opinion mining, social networks analysis, data mining, internet marketing, social media marketing, SMM, educational institution, educational organization, information model

В настоящее время многие организации заинтересованы в эффективной рекламе своих услуг и продуктов, наибольший интерес при этом представляет маркетинг в социальных сетях и блогосфере. В первую очередь это вызвано быстрорастущим процентом трафика социальных сетей в сети Интернет, возможностью таргетированного маркетинга, большим числом активных пользователей (число зарегистрированных пользователей социальной сети «Вконтакте» достигает 380 миллионов пользователей, более 80 миллионов посетителей заходят на сайт каждый день по официальным данным социальной сети «Вконтакте» на октябрь 2016 года), число пользователей Facebook превысило полтора миллиарда [11].

Образовательные организации также заинтересованы в рекламе своих услуг и конкурентных преимуществ в целях привлечения большего числа абитуриентов и партнеров. В работе [2] рассмотрена специфика рекламных коммуникаций образовательных учреждений с экономической точки зрения. При этом указывается, что при разработке рекламной кампании образовательным учреждениям необходимо придерживаться алгоритма, основанно-

го на особенностях поведения потребителя и учитывающего интересы всех субъектов взаимодействия. В работах [1, 6, 9] проблематика интернет-рекламы рассмотрена с социологической точки зрения: описаны особенности интернет-маркетинга в российском обществе, определены подходы к распространению рекламы, способы достижения социального эффекта и необходимость в обратной связи между рекламодателем и потребителем. В работах [13, 14] интернет-маркетинг рассматривается как один из самых эффективных инструментов интернет-бизнеса.

На основании обзора научной литературы можно выделить одно из перспективных направлений в области интернет-маркетинга – Social media marketing (SMM). SMM – это процесс привлечения трафика или внимания к бренду через социальные платформы. Он появился на рынке маркетинга менее десяти лет назад и успел доказать свою высокую эффективность, при этом все больше компаний разного уровня и специфики прибегают к его помощи. SMM является скрытой и ненавязчивой рекламой, позволяет познакомить целевую аудиторию с брендом, привлечь внимание к собственной деятельности, анонсировать и рекламировать товары и услуги, увеличивая посещаемость сайта.

Существует универсальная концепция SMM-стратегии [12]. Она подходит в том числе образовательным организациям, и в данной предметной области некоторые ее этапы могут быть заранее формализованы. SMM-стратегия образовательного учреждения представлена на рис. 1 (пунктирными линиями обозначены необязательные связи).

Применительно к образовательным организациям первоочередной долгосрочной целью маркетинга является привлечение внимания к своему бренду со стороны абитуриентов, что в конечном итоге должно привести к увеличению числа заявок на обучение.

Отдельно стоит отметить этап аналитики, поскольку он является одним из наиболее важных и в большей степени влияет на все остальные этапы. Основные задачи аналитики социальных сетей:

– мониторинг социальных сетей и блогов;

– проведение аналитики тональности упоминаний бренда;

– поиск источников негатива в социальных сетях и блогах;

– проведение исследований в социальных сетях и определение природы негатива;

– проведение аналитики информационных поводов (инфоповод; событие, служащее формированию и информационной поддержке сообщества, формированию иного взгляда, коррекции взгляда на предмет информационного повода);

– анализ эффективности рекламной компании;

– отслеживание источников и качества трафика;

– проведение аналитики изменения информационного поля.

Рис. 1. SMM-стратегия образовательного учреждения

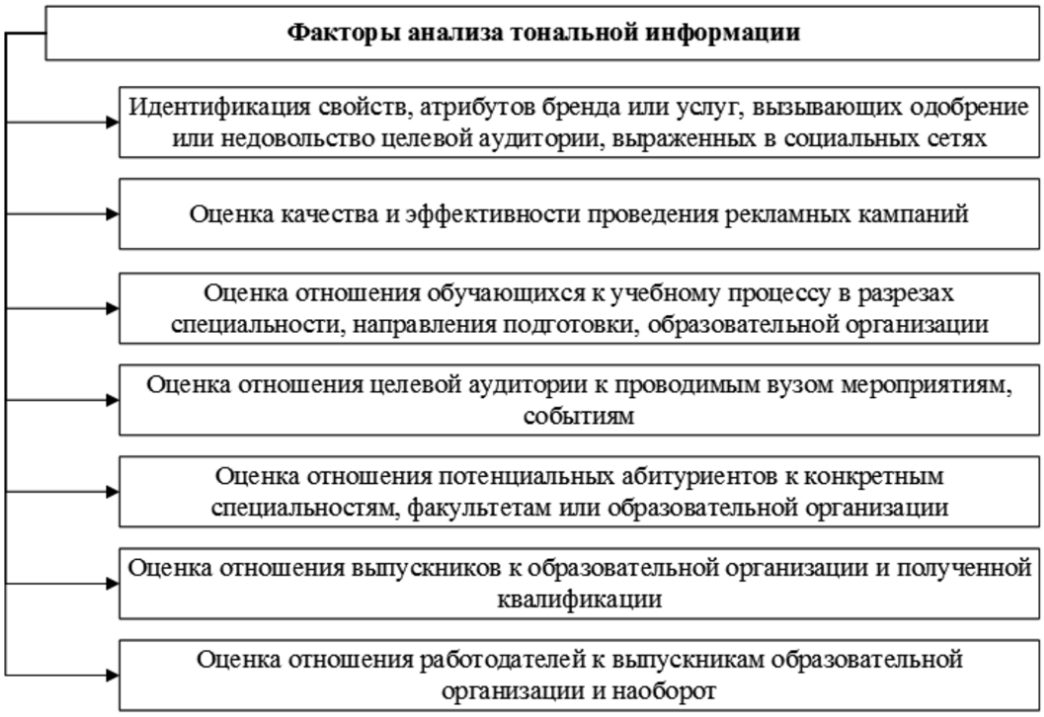

Рис. 2. Критерии анализа тональной информации

Для решения представленных выше задач необходимы не только квалифицированные специалисты, но и эффективные инструменты мониторинга, сбора и анализа информации из социальных сетей, а также инструментарий интеллектуального анализа мнений.

Особое место в области аналитики SMM отводится идентификации тональных сообщений социальных сетей (позитив / негатив / нейтрально). С помощью данной информации можно качественно оценить эффективность деятельности образовательного учреждения, выявить целый ряд факторов, влияющих на имидж бренда. Кроме того, полученная информация может быть использована руководством при принятии решений в процессе управления деятельностью образовательной организации. Актуальные критерии анализа тональной информации в предметной области образовательных учреждений представлены на рис. 2.

Поиск необходимой информации выполняется как вручную, так и с помощью технологий и программных средств мониторинга и извлечения информации из социальных сетей. Поиск и анализ негативных и позитивных мнений в извлеченной информации выполняется вручную. Однако, так как выражаемые пользователями социальных сетей мнения являются текстовыми, данную задачу с математической точки зрения можно свести к задаче анализа тональности текста (Sentiment Analysis).

Анализ тональности текста – класс математических методов обработки текстов на естественном языке в целях выявления и изучения эмоциональной составляющей текста. Задача анализа тональности является частной задачей классификации текстов и извлечения информации, которая лежит в области компьютерной (математической) лингвистики, возникшей на стыке таких наук, как лингвистика, математика, информатика (Computer Science) и искусственный интеллект [4, с. 2].

Так же в работе [4] описаны основные подходы к решению задачи анализа тональности текста, среди которых стоит отметить аспектный (интеллектуальный) подход. Основной идеей данного подхода является извлечение и анализ мнений (opinion mining), содержащих тональную оценку. В основе данного подхода лежат методы машинного обучения. Это обусловлено в первую очередь повышением количества и объемов текстовых корпусов, а также непрерывно увеличивающимся объемом неструктурированной информации в сети Интернет. Все множество методов машинного обучения можно разделить на две категории:

– машинное обучение без учителя (unsupervised machine learning);

– машинное обучение с учителем (supervised machine learning).

Главной особенностью методов машинного обучения без учителя является отсутствие необходимости в большой обучающей коллекции. Однако данные методы не являются часто используемыми при анализе русскоязычных текстов ввиду лингвистических особенностей русского языка и сложности реализации самих алгоритмов, что приводит к низким показателям точности.

Методы машинного обучения без учителя, в свою очередь, нацелены на работу с большими массивами данных, обладают хорошими показателями качества классификации, достаточно просты в применении, что обуславливает их широкое распространение. Данные методы основаны на предположении о том, что термины, которые чаще встречаются в анализируемом документе и в то же время присутствуют в небольшом количестве документов всей коллекции, имеют наибольший вес. Определив тональность данных терминов, можно сделать вывод о тональности всего документа [3]. Для методов машинного обучения с учителем необходимо предварительное составление и разметка тестовой коллекции, необходимой для обучения классификатора, затем формируется векторное представление каждого документа в рамках модели «bag of words» или «bag of features», для каждого документа вручную указывается тип тональности и производится обучение классификатора. Полученная модель используется для определения тональности документов новой коллекции. Рассмотрим основные методы машинного обучения, применяемые для решения задачи анализа тональности текста и проведем их сравнительный анализ:

1. Метод опорных векторов (Support Vector Machine, SVM).

SVM является одним из наиболее популярных методов машинного обучения с учителем. Основная идея метода заключается в переводе исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Разделяющая гиперплоскость строится на основе двух параллельных гиперплоскостей, разделяющих классы.

2. Наивный байесовский классификатор (Naive Bayes classifier).

Данный классификатор основан на применении теоремы Байеса со строгими (наивными) предположениями о независимости. Для объекта классификации вычисляются функции правдоподобия каждого из классов, по которым вычисляются апостериорные вероятности классов. Объект относится к тому классу, для которого апостериорная вероятность максимальна (принцип максимума апостериорной вероятности).

3. Метод k ближайших соседей (k-nearest neighbors algorithm, k-NN).

Метрический классификатор, основанный на оценивании сходства объектов – классифицируемый объект относится к тому классу, которому принадлежат ближайшие к нему объекты обучающей выборки. Помимо низкой точности, этот классификатор имеет проблему со скоростью классификации: если в обучающей выборке N объектов, в тестовой выборке M объектов, а размерность пространства – K, то количество операций O для классификации тестовой выборки может быть оценено как

$$O = K*M*N.$$

Однако, несмотря на описанные недостатки, данный метод очень прост в применении и полезен для апробации применяемых при анализе лингвистических технологий.

Для проведения сравнительного анализа методов для задачи анализа тональности принято использовать стандартные метрики оценки качества текстовой классификации – точность (precision), полнота (recall), F-метрика (f-measure, F1). В исследовании [10] был произведен эксперимент по оценке качества методов машинного обучения: SVM – классификатор метода опорных векторов, GMM – байесовский классификатор на основе смеси многомерных нормальных распределений, ROC – классификатор Роччио, KNN – классификатор к-ближайших соседей, VMF – классификатор фон Мизеса – Фишера, TREE – классификатор на основе деревьев решений. Метрики точности и полноты рассматривались отдельно для позитивных и негативных результатов. Результаты анализа представлены в таблице.

Сравнительный анализ методов машинного обучения без учителя

Method	P_p	R_p	P_n	R_n
SVM	0,86	0,99	0,41	0,44
GMM	0,88	0,73	0,27	0,42
ROC	0,92	0,18	0,27	0,8
KNN	0,87	0,78	0,23	0,30
VMF	0,94	0,47	0,31	0,57
TREE	0,90	0,70	0,27	0,30

Как можно заметить из приведенной таблицы, показатели качества для отрицательных текстов заметно хуже, чем для положительных. При этом наиболее высокие

показатели продемонстрировали алгоритмы: SVM, KNN, TREE. SVM имеет, как правило, высокую точность, но низкую полноту, что может давать не слишком хорошую метрику F1. Для всех методов свойственно низкое значение показателя полноты вследствие плохого распознавания отрицательного лексикона.

Стоит отметить, что математический аппарат методов машинного обучения за последние годы не претерпел значительных изменений, поэтому повышение качества классификации достигается в основном за счет применения лингвистических технологий анализа. Приведенные данные показывают, что для задачи анализа тональности мнений пользователей социальных сетей целесообразным является применение метода SVM или наивного байесовского классификатора, поскольку они обладают достаточно хорошими показателями качества.

На основе вышеизложенного можно построить информационную модель управления деятельностью университета с использованием технологий SMM, анализа социальных сетей и анализа тональности мнений. Модель представлена на рис. 3.

В данной модели выделены три основных процесса:

– мониторинг общественного мнения, выражаемого пользователями социальных сетей;

– системный анализ мнений;

– разработка и применение SMM-стратегии на основе обратной связи, реализованной предыдущими процессами.

Представленная модель может быть интерпретирована как модель коммерческих рынков b2c и b2b, описанных в [2]. Рынок b2c (business-to-consumer) – коммерческие взаимоотношения между организацией (business) и частным (конечным) потребителем (consumer). Применимо к описанной задаче в качестве b2c-рынка выступает рынок образовательных услуг. Рынок b2b (business-to-business) – вид информационного и экономического взаимодействия, субъектами которого выступают юридические лица (компании), которые работают не на конечного рядового потребителя, а на такие же компании, то есть на другой бизнес. В рамках проводимого исследования в качестве b2b-рынка выступает рынок труда.

Описанная информационная модель позволяет наглядно увидеть взаимосвязи этих рынков и целевой аудитории SMM-стратегии, решает проблему нехватки обратной связи о потребностях специалистов того или иного профиля. Таким образом может быть решена острая проблема рассогласования рынка образовательных услуг и рынка труда, приводящая к уменьшению эффективности использования человеческих ресурсов и снижению уровня социально-экономического развития на региональном уровне [2].

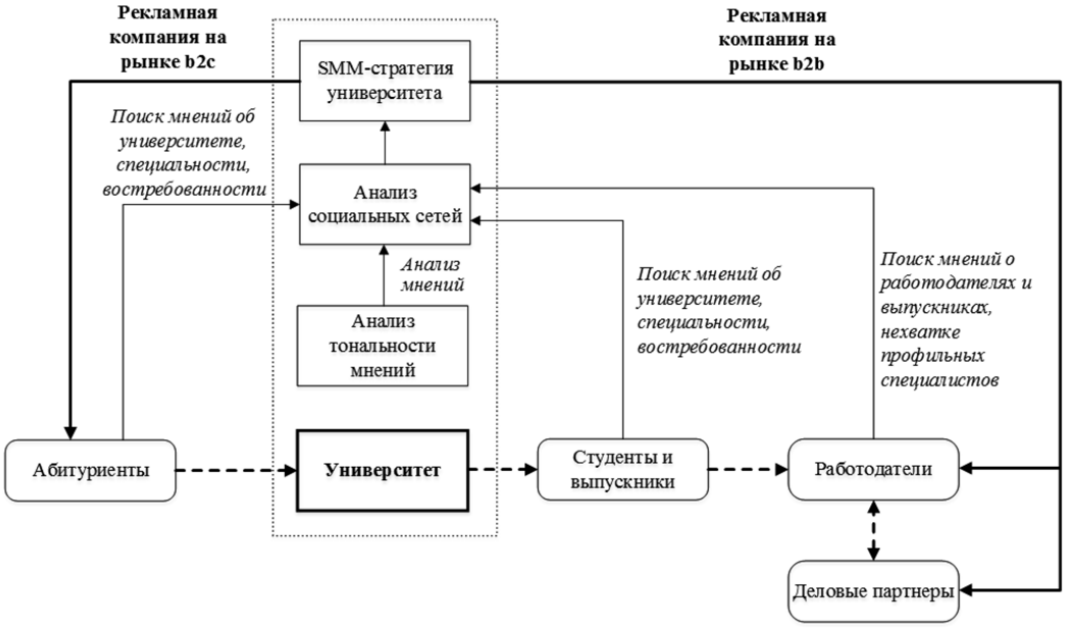

Рис. 3. Информационная модель управления деятельностью университета с использованием технологий SMM и анализа социальных сетей

Поскольку технологии анализа тональности могут быть использованы для анализа отношения абитуриентов и учащихся к специальностям и направлениям подготовки, то получаемая информация может быть использована не только в целях проведения рекламной кампании, но и в процессе оптимизации и управления деятельностью образовательной организации. К примеру, эта информация может быть использована университетом для решения задачи корректировки числа планируемых бюджетных мест для обучения.

Стоит отметить, что технологии анализа социальных сетей и анализа тональности мнений представляют отдельные, активно развивающиеся области научных исследований [7, 8]. Также в настоящее время инструменты мониторинга социальных сетей плохо автоматизированы и требуют большого человеческого участия, а выполняемый в дальнейшем анализ проводится вручную [3]. Полноценные интеллектуальные инструменты мониторинга и анализа мнений находятся в начальной стадии развития и представляют большой научный и коммерческий интерес [5].

Список литературы

1. Бочарова Е.В. Коммерческая интернет-реклама в российском обществе: социальная среда формирования и социальная эффективность: Автореф. дис. канд. соц. наук. Ростов-на-Дону: ЮФУ, 2012. – 35 с.

2. Витальева Е.М. Специфика рекламных коммуникаций учреждений высшего профессионального образования на рынке труда: Автореф. дис. канд. экон. наук. – Волгоград: ВолгГТУ, 2014. – 24 с.

3. Гавриков М.М. Теоретические основы разработки и реализации языков программирования: учебное пособие для студентов вузов, обучающихся по специальности «Программное обеспечение вычислительной техники и автоматизированных систем» направления подготовки дипломированных специалистов «Информатика и вычислительная техника» / Иванченко А.Н., Гринченков Д.В., Москва, 2010. – 184 с.

4. Горбушин Д.А. Системный анализ подходов к решению задачи идентификации тональности текста / Гринченков Д.В., Мохов В.А., Нгуен Ф.Х. // Известия высших учебных заведений. Северо-Кавказский регион. Серия: Технические науки. – 2016. – № 2 (190). – С. 36–41.

5. Горбушин Д.А. Анализ методов автоматической классификации тональности текста // Научно-техническая конференция инновационных проектов: сб. материалов конф., г. Новочеркасск, 14–16 дек. 2014 г. / Юж.-Рос. гос. политехн. ун-т им. М.И. Платова – Новочеркасск: Лик, 2014. – С. 123–125.

6. Гордеев А.И. Интернет-реклама в коммуникационном пространстве российского общества: функции, специфика и перспективы развития: Автореф. дис. канд. соц. наук. Ростов-на-Дону: ЮФУ, 2013. – 29 с.

7. Гринченков Д.В., Кущий Д.Н. Об одном подходе к оценке информационной значимости результатов поиска путем семантической оптимизации запроса // Всероссийская научно-практическая конференция Информационно-телекоммуникационные системы и технологии. Новочеркасск, 2015. – С. 74.

8. Кущий Д.Н., Гринченков Д.В. К вопросу о многоуровневой аналитической обработке текстов на естественном языке // Всероссийская научно-практическая конференция Информационно-телекоммуникационные системы и технологии. – Новочеркасск, 2015. – С. 176.

9. Лысенкова М.Ф. Политическая интернет-реклама как разновидность избирательных технологий в современной России: особенности и тенденции развития: Автореф. дис. канд. полит. наук. – Москва: МосГУ, 2013. – 26 с.

10. Худякова М.В., Васильев В.Г., Давыдов С. Классификация отзывов пользователей с использованием фрагментных правил // Компьютерная лингвистика и интеллектуальные технологии: по материалам ежегодной международной конференции «Диалог». – 2012. – № 11 (18). – С. 66–76.

11. Число пользователей Facebook превысило полтора миллиарда [Электронный ресурс]. – Режим доступа: https://lenta.ru/news/2015/11/05/facebook_accounts/ (дата обращения: 01.10.2016).

12. Ashley C., Tuten T. Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement // Psychology & Marketing. – 2015. – Т. 32. – № 1. – P. 15–27.

13. Goi C.L. Marketing Mix: A review of ‘P’ // The Journal of Internet Banking and Commerce. – 2015. – P. 2005.

14. Leung X.Y., Xue L., Bai B. Internet marketing research in hospitality and tourism: a review and journal preferences // International Journal of Contemporary Hospitality Management. – 2015. – Т. 27. – № 7. – P. 1556–1572.

References

1. Bocharova E.V. Kommercheskaja internet-reklama v rossijskom obshhestve: socialnaja sreda formirovanija i socialnaja jeffektivnost: Avtoref. dis. kand. soc. nauk. Rostov-na-Donu: JuFU, 2012. 35 p.

2. Vitaleva E.M. Specifika reklamnyh kommunikacij uchrezhdenij vysshego professionalnogo obrazovanija na rynke truda: Avtoref. dis. kand. jekon. nauk. Volgograd: VolgGTU, 2014. 24 p.

3. Gavrikov M.M. Teoreticheskie osnovy razrabotki i realizacii jazykov programmirovanija: uchebnoe posobie dlja studentov vuzov, obuchajushhihsja po specialnosti «Programmnoe obespechenie vychislitelnoj tehniki i avtomatizirovannyh sistem» napravlenija podgotovki diplomirovannyh specialistov «Informatika i vychislitelnaja tehnika» / Ivanchenko A.N., Grinchenkov D.V., Moskva, 2010. 184 p.

4. Gorbushin D.A. Sistemnyj analiz podhodov k resheniju zadachi identifikacii tonalnosti teksta / Grinchenkov D.V., Mohov V.A., Nguen F.H. // Izvestija vysshih uchebnyh zavedenij. Severo-Kavkazskij region. Serija: Tehnicheskie nauki. 2016. no. 2 (190). pp. 36–41.

5. Gorbushin D.A. Analiz metodov avtomaticheskoj klassifikacii tonalnosti teksta // Nauchno-tehnicheskaja konferencija innovacionnyh proektov: sb. materialov konf., g. Novocherkassk, 14–16 dek. 2014 g. / Juzh.-Ros. gos. politehn. un-t im. M.I. Platova Novocherkassk: Lik, 2014. pp. 123–125.

6. Gordeev A.I. Internet-reklama v kommunikacionnom prostranstve rossijskogo obshhestva: funkcii, specifika i perspektivy razvitija: Avtoref. dis. kand. soc. nauk. Rostov-na-Donu: JuFU, 2013. 29 p.

7. Grinchenkov D.V., Kushhij D.N. Ob odnom podhode k ocenke informacionnoj znachimosti rezultatov poiska putem semanticheskoj optimizacii zaprosa // Vserossijskaja nauchno-prakticheskaja konferencija Informacionno-telekommunikacionnye sistemy i tehnologii. Novocherkassk, 2015. pp. 74.

8. Kushhij D.N., Grinchenkov D.V. K voprosu o mnogourovnevoj analiticheskoj obrabotke tekstov na estestvennom jazyke // Vserossijskaja nauchno-prakticheskaja konferencija Informacionno-telekommunikacionnye sistemy i tehnologii. Novocherkassk, 2015. pp. 176.

9. Lysenkova M.F. Politicheskaja internet-reklama kak raznovidnost izbiratelnyh tehnologij v sovremennoj Rossii: osobennosti i tendencii razvitija: Avtoref. dis. kand. polit. nauk. Moskva: MosGU, 2013. 26 p.

10. Hudjakova M.V., Vasilev V.G., Davydov S. Klassifikacija otzyvov polzovatelej s ispolzovaniem fragmentnyh pravil // Kompjuternaja lingvistika i intellektualnye tehnologii: po materialam ezhegodnoj mezhdunarodnoj konferencii «Dialog». 2012. no. 11 (18). pp. 66–76.

11. Chislo polzovatelej Facebook prevysilo poltora milliarda [Jelektronnyj resurs]. Rezhim dostupa: https://lenta.ru/news/2015/11/05/facebook_accounts/ (data obrashhenija: 01.10.2016).

12. Ashley C., Tuten T. Creative strategies in social media marketing: An exploratory study of branded social content and consumer engagement // Psychology & Marketing. 2015. T. 32. no. 1. pp. 15–27.

13. Goi C.L. Marketing Mix: A review of P // The Journal of Internet Banking and Commerce. 2015. pp. 2005.

14. Leung X.Y., Xue L., Bai B. Internet marketing research in hospitality and tourism: a review and journal preferences // International Journal of Contemporary Hospitality Management. 2015. T. 27. no. 7. pp. 1556–1572.