

# Проблемы информационного пространства и информационной безопасности

УДК 004.4

С.А. Малышко

## Мониторинг и анализ медиaproстранства – информационно-аналитическая система *NEWSCAPE*

Описаны основные возможности информационно-аналитической системы *NEWSCAPE* – эффективной информационной технологии мониторинга информационного пространства и основанный на ней онлайн-сервис. *NEWSCAPE* содержит различные математические методы и алгоритмы, ориентированные на мониторинг и многофакторный анализ медиaproстранства (информационного пространства).

The main features of the information-analytical system *NEWSCAPE* are described, which is an efficient information technology of monitoring the media space and online service based on it. *NEWSCAPE* accumulates a variety of the mathematical methods and the algorithms focused on monitoring and multivariate analysis of media space (information space).

Описано основні можливості інформаційно-аналітичної системи *NEWSCAPE*, яка є ефективною інформаційною технологією, та заснований на ній онлайн-сервіс. *NEWSCAPE* містить різні математичні методи і алгоритми, орієнтовані на моніторинг та багатофакторний аналіз медиaproстор (інформаційного простору).

**Введение.** Бурное развитие глобальных сетей электронных коммуникаций, базирующееся на всеобщей компьютеризации, привело к появлению огромного медиaproстранства. Всего несколько десятилетий тому назад актуальной была задача по поиску нужной информации, в то время как теперь основной проблемой становится отсеивание ненужных данных [1]. Громадные объемы неструктурированной информации обрушиваются на пользователя при любом поисковом запросе. На первый план при анализе информационного пространства выходят задачи поиска закономерностей и взаимосвязей, отслеживания тенденций и трендов, прогнозирование роста и падения числовых показателей объектов в медиaproстранстве. Информационно-аналитические системы позволяют увидеть «простое и понятное» в огромном, сложном и хаотическом информационном океане.

Актуальность и практическая востребованность исследуемой в статье проблематики растет вместе с расширением и количественным увеличением объема медиaproстранства. Сегодня не существует некой единой стандартизированной технологии мониторинга и анализа информационного пространства. Основная цель

разработки описываемой информационно-аналитической системы (ИАС) *NEWSCAPE* – создание современной эффективной информационной технологии и основанного на ней онлайн-сервиса, аккумулирующих различные математические методы и алгоритмы, ориентированные на мониторинг и многофакторный анализ медиaproстранства. Некоторые идеи, методы и алгоритмы, используемые в ИАС, – это оригинальные авторские разработки [2]. Обзор существующих систем мониторинга информационного пространства выходит за рамки данной публикации вследствие большого разнообразия и специфики программных продуктов такого рода. В качестве примера таких систем отметим актуальную технологию интеграции новостей *InfoStream* [3, 4]. В работе обсуждаются вопросы разработки и использования ИАС *NEWSCAPE*. Как и многие из существующих систем указанного назначения, ИАС *NEWSCAPE* продолжает активно развиваться и совершенствоваться в соответствии с потребностями практики и современными требованиями к эффективному программному обеспечению, ориентированному на широкий круг пользователей. Расширяется также и аналитический инструментальный системы, осно-

ванный на современных математических методах и алгоритмах.

Созданная ИАС базируется на современных научно-технических разработках сотрудников Института кибернетики им. В.М. Глушкова НАН Украины в сфере математического моделирования и ИТ, развитых и реализованных усилиями объединенного коллектива высококвалифицированных специалистов, и представляет собой современный программно-аппаратный комплекс, реализующий основные функции мониторинга медиапространства.

### **Общие сведения о системе**

ИАС *NEWSCAPE* ориентирована на использование в автономном режиме и в качестве интеллектуальной компоненты системы поддержки принятия решений (СППР) в трудноформализуемых и конфликтных ситуациях. Разработанная система, развитие которой продолжается, обеспечивает процесс выработки эффективных решений различного уровня: от задач оперативного функционирования ситуационного центра до проблем стратегического планирования. ИАС также может быть эффективно использована и в современной информационной войне – как для контроля и планирования действий «своих», так и для выявления и анализа действий «чужих».

Реализованные программно-алгоритмические средства ИАС позволяют:

- обеспечить непрерывный автоматизированный мониторинг медиапространства, включающего в себя около тысячи наиболее рейтинговых сайтов, основные государственные и региональные интернет-источники, ведущие печатные СМИ и телевизионные каналы (видеомониторинг);
- аккумулировать разнообразные разрозненные базы данных в единую информационную систему;
- накапливать в структурированном виде результаты мониторинга для последующего анализа;
- использовать современные методы количественного и качественного анализа информации;
- обеспечить эффективную коллективную работу операторов и аналитиков разного уровня.

Процесс анализа медиапространства с использованием рассматриваемой ИАС основан на анализе категорий, формируемых пользователями и характеризующих исследуемые структурные единицы – объекты, персоны, процессы, события и пр.

Архитектура и функциональное наполнение ИАС создавалось с целью достижения таких целей, как обеспечение возможности:

- оперативно получать необходимую и достоверную информацию;
- отслеживать основные тенденции медиапространства;
- проводить многофакторный анализ объектов, персон, событий, тенденций, трендов;
- совершенствовать управление масштабными системами и подсистемами;
- моделировать и исследовать сложные процессы;
- прогнозировать возможные сценарии развития процессов.

Предлагаемая технология создания ИАС базируется на основе использования современных веб-приложений, что обеспечивает оперативность и доступность ИАС в любых точках, имеющих подключение к интернету, с возможностью использования терминальных устройств различного типа – компьютеры, планшеты, смартфоны.

ИАС *NEWSCAPE* может быть эффективно использована в органах государственной власти (Верховная Рада и ее комитеты, Администрация Президента, Кабинет Министров Украины), в крупных компаниях, организациях, учреждениях, политических партиях как система анализа и поддержки принятия решений на основе экономической и социально-политической информации. Системы подобного класса востребованы в различных медиа-холдингах, крупных корпорациях и объединениях, активно продвигающих свои товары или услуги, что находит отражение в информационном пространстве.

### **Основные возможности системы**

На основе разработанных оригинальных математических методов и алгоритмов анализа

пользователям доступен развитый инструментарий для анализа информации, обеспечивающий:

- анализ характеристик объектов в информационном пространстве, в частности, таких, как количество упоминаний, индекс цитирования, индекс медиаактивности, индекс информационной активности, рейтинг информационной динамики, региональное распределение, распределение по СМИ, оценка уровня агрессивности текстов, выявление манипуляторных признаков в тексте;

- контент-анализ – количественный, частотный, разностный, а также анализ соседних слов и частей речи в предложениях;

- категорийный анализ, объектами которого могут выступать политики, политологи, политические партии, области Украины, страны мира, столицы государств, международные организации, промышленность, продукты питания и др.;

- морфологический анализ, оперирующий со существительными, местоимениями, прилагательными и наречиями, глагольными формами, несловарными объектами, другими частями речи;

- анализ количественных характеристик текстов – предложений, слов, существительных, уникальных существительных, лемм, словоформ с возможностью расчета ряда важнейших статистических показателей.

### Мониторинг информационного пространства

Существуют различные подходы к определению термина *информационное пространство*. На практике это понятие часто относят к общеупотребляемым терминам, не требующим дополнительных объяснений. Однако многие исследователи считают, что современное значение понятия *информационное пространство* сложилось в результате эволюции концептуальной схемы различения в совокупном геополитическом пространстве областей, обладающих свойствами, по-

зволяющими рассматривать их как самостоятельные пространства со своими границами, структурой, ресурсами и особенностями взаимодействия субъектов социальных отношений [5]. С повышением уровня компьютеризации общества вместо термина *информационное пространство* используется современный синоним – медиапространство, т.е. пространство, создаваемое электронными средствами коммуникации. Для наших целей под информационным пространством будем понимать совокупность электронных версий традиционных СМИ, онлайн-ресурсы (сайты, порталы) и «новые медиа» (социальные сети, блоги).

Базовая функция ИАС *NEWSCAPE* – это непрерывный автоматизированный мониторинг медиапространства, включающего в себя около тысячи наиболее рейтинговых сайтов, основные государственные и региональные интернет-источники, ведущие печатные СМИ и телевизионные каналы (видеомониторинг). Развитая специализированная поисковая подсистема обеспечивает пользователю возможность формировать сложные запросы к ИАС для получения результатов мониторинга по интересующему информационному объекту.

Рассмотрим возможности поисковой подсистемы. Основной синтаксис текстовых запросов представлен на рис. 1.

Первичные числовые результаты выполнения поискового запроса с временным графиком, отображающим количество найденных релевантных новостей и индекс цитирования, представлены на рис. 2.

В качестве иллюстрации возможностей системы был выбран виртуальный объект *Кличко*

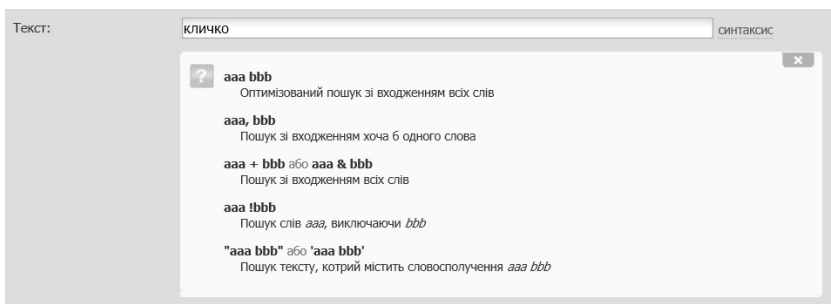


Рис. 1

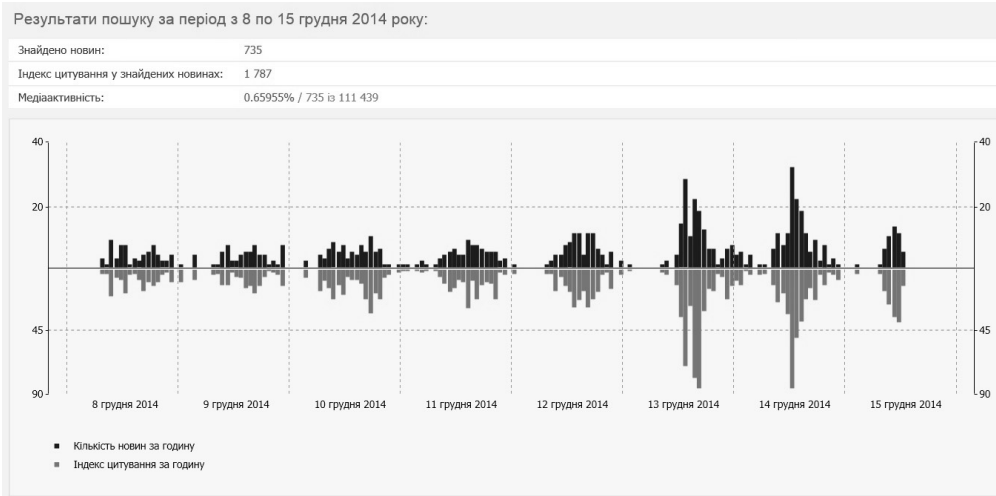


Рис. 2

Розподіл пошукових слів

Сортування по: **кількості новин** | індексу цитування | Показники: абсолютні | відносні

	8 грудня 2014	9 грудня 2014	10 грудня 2014	11 грудня 2014	12 грудня 2014	13 грудня 2014	14 грудня 2014	15 грудня 2014	Загалом
кличко	64   129	63   136	82   223	81   181	92   209	138   410	160   369	47   130	727   1 787

Медиаактивність

Сортування по: абсолютні | відносні

	8 грудня 2014	9 грудня 2014	10 грудня 2014	11 грудня 2014	12 грудня 2014	13 грудня 2014	14 грудня 2014	15 грудня 2014	Загалом
Сліввідношення	64 із 17 557	63 із 17 769	83 із 18 497	82 із 19 029	93 із 17 935	141 із 7 297	162 із 6 360	47 із 6 995	735 із 111 439

Рис. 3

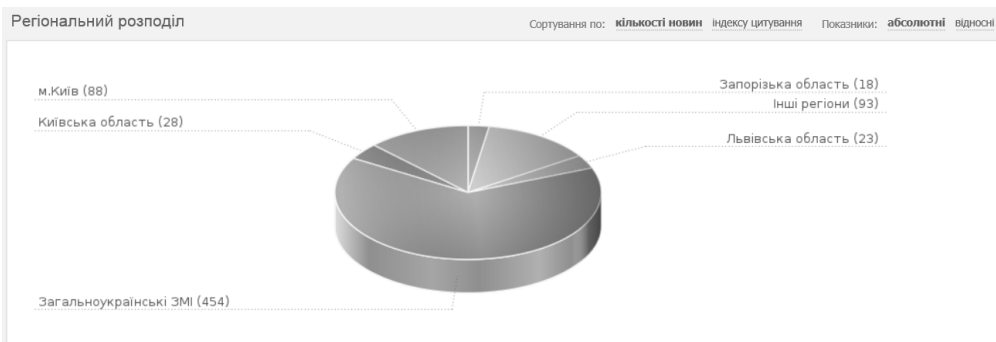


Рис. 4

	8 грудня 2014	9 грудня 2014	10 грудня 2014	11 грудня 2014	12 грудня 2014	13 грудня 2014	14 грудня 2014	15 грудня 2014
Загальноукраїнські ЗМІ	29   45	36   77	52   153	53   120	54   122	99   278	104   246	27   79
м.Київ	16   40	14   29	9   21	8   10	12   25	8   20	8   23	13   30
Київська область	2   1	7   2	3   3	6   3	16   7	22   6	17   6	0   0
Львівська область	0   0	4   1	1   2	6   2	7   3	3   4	19   10	2   1
Запорізька область	9   2	4   1	1   1	6   2	4   3	16   3	10   5	2   1
інші регіони	14   27	5   6	11   24	12   28	9   20	15   38	23   45	4   14
	62   123	59   127	78   203	80   176	88   194	135   377	156   360	46   127

Рис. 5

и период мониторинга с 8/12/14 по 15/12/14. Все числовые показатели – реальные результаты мониторинга медиaprостранства Украины

за указанный период. Введем следующие обозначения и продемонстрируем их значения на тестовом запросе по объекту:

$S$  – общее количество информационных блоков в системе за период (111439);

$N$  – количество найденных информационных блоков по запросу *Кличко* (735);

$K$  – индекс цитирования, равный количеству найденных слов *Кличко* (1787);

$M$  – показатель медиаактивности, равный  $N / S * 100$  процентов (0,69 процента);

$L$  – индекс медиаприсутствия объекта, равный  $K / N$  (2,43).

Дополнительно система позволяет получить базовые числовые характеристики с точностью до дня (рис. 3).

Отражение этих показателей в региональном разрезе по областям Украины приведено на рис. 4, 5.

Для углубленного анализа часто необходимо иметь сведения об основных источниках информации. На рис. 6 приведены сведения для подобного анализа.

Очевидно, для серьезного исследования объектов в информационном пространстве необходимы дополнительные действия – в

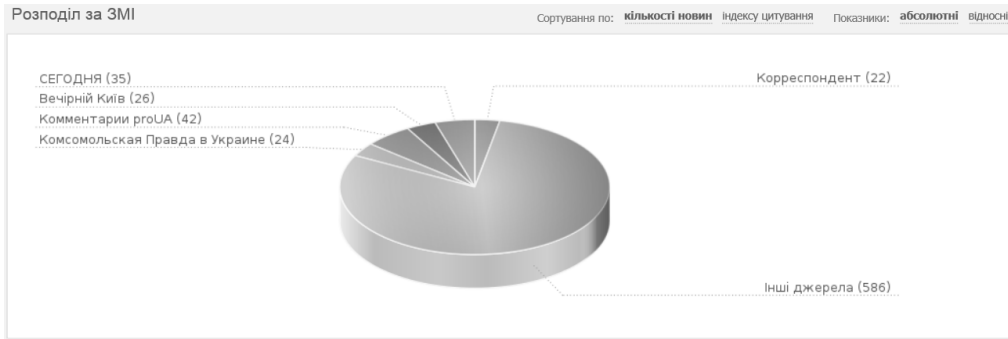


Рис. 6

24. Кличко хочет снести Дом Профсоюзов – The Kiev Times / 14.12.2014 17:26 #28717109

Сгоревшее здание на Майдане восстанавливать опасно Мэр Киева Виталий **Кличко** заявил о необходимости сноса Дома профсоюзов в Киеве из-за значительных поврежденных конструкции. Об этом он заявил в интервью «Радио Свобода». «Сейчас мы это решаем. Есть предварительные исследования, согласно которым этот дом настолько поврежден, что его нужно сносить и строить новый. То есть конструкции...

детальніше

---

25. Будинок профспілок можуть знести - Кличко – Коментарии proUA / 15.12.2014 09:00 #28748526

Столичный Будинок профспілок настільки постраждав від пожежі, що відновлювати його небезпечно Про це заявив мер Києва Віталій **Кличко**, передає «Українська правда». Відповідаючи на питання, що буде з Будинком профспілок, міський голова зазначив, що наразі це питання вирішується. «Є попередні дослідження, згідно з якими цей будинок настільки пошкоджений, що його потрібно зносити і будувати...

детальніше

← 1 2 3 4 5 6 7 8 9 10 11 12 13 ... 31 →

зберегти результат пошуку

Рис. 7

11. Кличко намерен вернуть эвакуаторы – Дело / 15.12.2014 10:46 #28752831

В следующем месяце столичная власть начнет проведение тендеров на обустройство в Киеве современных парковок. Об этом глава Киевской городской государственной администрации Виталий **Кличко** сообщил в интервью "Радио Свобода".

По его словам, сейчас в столице насчитывается около 1 млн автомобилей. При этом парковочных мест в Киеве — только 6 тыс.

"Этих мест должно быть намного больше. Но мы должны создавать надлежащие условия для парковки автомобилей, чтобы не было случаев незаконной парковки на Крещатике, пешеходных зонах и тому подобное. Поэтому со следующего месяца мы начнем публичные аукционы и тендеры на обустройство в столице современных парковок", — сказал мэр столицы.

В то же время **Кличко** отметил, что для предотвращения незаконной парковки, которая блокирует и мешает движению автотранспорта, планируется привлечь эвакуаторы.

згорнути

новина від: 15.12.2014 10:46  
джерело: Дело / Друківані ЗМІ – центральні адреса новини: <http://delo.ua/ukraine/klichko-nameren-vozvrashchat-evakuary-285888/>  
рубрика за джерелом: Киев

Індекс цитування слів

слово	індекс
кличко	2

Рис. 8

рассматриваемом примере уточняющие запросы *Владимир Кличко, Виталий Кличко, Кличко + мэр* и другие позволят структурировать множество информационных блоков и исследовать взаимосвязи объектов мониторинга. Выделяющийся на графике (см. рис. 2) «всплеск» информационной активности за 13–14 декабря

требует более внимательного изучения с целью определения перечня событий, приведших к заметному росту информационного присутствия исследуемых объектов.

### Анализ результатов поиска

ИАС *NEWSCAPE* предоставляет пользователям удобный интерфейс для быстрой навигации по отобранному в результате поискового запроса информационным блокам (рис. 7). В сформированном ИАС-тексте выделены совпадения с объектом поиска, при этом поддерживается краткая и развернутая формы отображения информации, дополнительные характеристики источника данных (рис. 8). Аналитику предоставляется возможность быстро провести содержательный анализ отобранной по запросу информации, просмотреть более детально основные публикации по объекту.

Для выявления «тонких» особенностей исследуемых объектов в медиапространстве требуется дополнительная обработка с помощью специализированных подсистем ИАС. Пользователь может сохранить результаты поиска как в виде финального отчета (формат *RTF*), так и в виде специализированного файла для передачи данных

в подсистему контент-анализа текстов. Подробное описание дополнительных подсистем ИАС *NEWSCAPE* выходит за рамки данной публикации.

**Заключение.** В настоящее время накоплен огромный объем информации, а также разрабатывается и реализуется ряд специализированных программных инструментов для работы с ней. В то же время следует отметить, что сегодня в Украине отсутствуют прикладные системы, охватывающие такой широкий круг проблем анализа медиапространства и имеющие такой развитый программно-алгоритмический инструментарий, как в ИАС *NEWSCAPE*.

Начиная с 2010 г., в системе накоплено более 14 млн. информационных блоков, более 100 тыс. статей печатных СМИ и тысячи новостийных видеосюжетов. Также в системе доступна структурированная по областям и районам Украины экономическая, социальная и политическая информация. Используются в системе и агрегированные данные по юридическим объектам, предприятиям и физическим персонам.

В дальнейшем планируется развитие системы в следующих направлениях: агрегирование большего количества данных из разнотипных источников информации (мнения экспертов, телевидение, радио, социальные сети, реклама), стандартизация и подключение дополни-

тельных баз данных и знаний, совершенствование существующего и разработка нового математического аппарата, аналитического инструментария и информационных технологий анализа больших данных. Это позволит использовать сложные математические модели и методы для повышения степени адекватности формируемых оценок и прогнозов, а также для формирования оперативных и качественных аналитических отчетов для лиц, принимающих ответственные решения.

1. *Майер-Шенбергер В., Кукьер К.* Большие данные. Революция, которая изменит то, как мы живем, работаем и мыслим, 2014. – М.: Манн, Иванов и Фербер, 2013. – 240 с.
2. *Гуляницкий Л.Ф., Сергиенко И.В., Малышко С.А.* О программных средствах поддержки принятия решений в задачах группового выбора // УСиМ. – 1993. – № 5. – С. 90–97.
3. *Ландэ Д.В.* Поиск знаний в *Internet*. – М.: Диалектика, 2005. – 272 с.
4. *Ландэ Д.В., Снарский А.А., Безсуднов И.В.* Интернетика: Навигация в сложных сетях: модели и алгоритмы. – М.: Либроком, 2009. – 264 с.
5. *Манойло А.В.* Государственная информационная политика в особых условиях. – М.: МИФИ, 2003. – 388 с.

Поступила 13.01.2015  
Тел. для справок: +38 044 526-5585 (Киев)  
E-mail: [smsmsm@admpr.com.ua](mailto:smsmsm@admpr.com.ua)  
© С.А. Малышко, 2015

## Внимание !

**Оформление подписки обязательно для желающих  
опубликовать статьи в нашем журнале.**

**В розничную продажу журнал не поступает.**

Подписной индекс 71008