

Востребованность автоматизированных инструментов при проведении исследований в гуманитарных науках*

Л.А. Бершадская, А.С. Биккулов

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики
bershadskaya.lyudmila@gmail.com, bikkulov@mail.ifmo.ru

Аннотация

В статье отражены основные результаты проведения комплексного исследования востребованности и использования автоматизированных инструментов в гуманитарных науках. В статье обосновывается актуальность и востребованность разработки веб-ориентированного научно-производственного центра в области социодинамики, ставятся актуальные научные проблемы для исследования.

1. Введение

На сегодняшний день в среде гуманитарных наук происходит сдвиг к исследованию социальных сетей всевозможных масштабов, и в то же время имеет место тенденция усиленного использования математического аппарата и автоматизированных средств сбора, обработки и анализа получаемых данных. Данный тезис касается практически всех областей современной гуманитарной науки.

О востребованности автоматизированных средств для проведения исследований говорят результаты исследования, включающего в себя:

- анализ статистики русскоязычных и англоязычных поисковых запросов по тематике измерений, анализа и моделирования процессов в социальных сетях;
- обзор российских и зарубежных исследовательских центров в сфере измерений, анализа и моделирования процессов в социальных сетях;
- анализ российских и зарубежных курсов/модулей обучения по социальному моделированию, применению компьютерных технологий в социологии, анализу социальных сетей.

Исследование было проведено в рамках проекта, реализованного в НИУ ИТМО «Разработка веб-ориентированного производственно-исследовательского центра в области социодинамики и ее приложений» (заказчик Минобрнауки РФ). Цель проведения исследования заключалась в обоснова-

нии востребованности разрабатываемого продукта на рынке гуманитарных исследований.

2. Анализ статистики поисковых запросов

О востребованности обозначенной тематики можно судить, в том числе по статистике поисковых запросов в популярных системах Яндекс и Google.

Анализ поисковых запросов охватил как русскоязычный, так и англоязычный сегменты Интернета. При анализе использовалось тематическое деление на следующие смысловые блоки поисковых запросов:

- проведение социологических исследований;
- средства и методы сбора и обработки информации;
- контент-анализ данных;
- методы математического моделирования социальных процессов.

Анализ поисковых запросов в системе Яндекс проводился по всем странам и регионам, но при этом русскоязычные запросы, естественно, существенно лидировали на территории России и стран СНГ, что позволяет относить сделанные на основе анализа этих поисковых запросов выводы именно к России и странам СНГ. Для получения более широкой международной картины был проведен анализ статистики англоязычных запросов в данной тематике в поисковой системе Google, что позволило выявить страны, с наибольшей поисковой активностью по исследуемым тематикам. Рассмотрим более подробно результаты анализа русскоязычных запросов.

По данным сервиса Яндекс-метрика, интерес пользователей этой поисковой системы относительно методов проведения социологических исследований социальных сетей достигает в среднем около 11 тысяч запросов в месяц. Чаще всего поисковые запросы этого типа принадлежат интернет-пользователям из Москвы, Санкт-Петербурга, Свердловской и Ростовской областей. Свыше 75 тысяч запросов в месяц приходится на запросы, связанные с обработкой информации и связанными с этим технологиями. Больше всего запросов

такого типа приходится на Москву, Республики Башкортостан и Дагестан, Ростовскую область и Санкт-Петербург.

Контент-анализ является одной из ключевых методик, применяющихся в традиционных социологических исследованиях блогосферы. Количество запросов по обобщенной тематике «контент-анализ» достигает свыше 21 тыс. в месяц. Чаще всего методикой контент-анализа интересуются пользователи из Санкт-Петербурга, Москвы, Свердловской, Новосибирской, Ростовской областей. Число поисковых запросов, касающихся непосредственно моделирования социальных процессов, достигает более 1,5 тыс. в месяц, большая доля этих запросов приходится на пользователей из Санкт-Петербурга, Челябинской, Самарской, Нижегородской областей и Приморского края.

Анализ статистики поисковых запросов позволил сделать следующие выводы:

1. У пользователей сети Интернет имеется устойчивый интерес к данной тематике: разброс количества поисковых запросов по тематике проведения социологических исследований, методов сбора и обработки данных, моделирования социальных процессов находится в диапазоне от 1,5 до 21 тыс. запросов в месяц, при этом примерно по 500 пользователей в месяц интересуются именно исследованиями и анализом социальных сетей. Считается, что 500-1000 целевых запросов в месяц уже отражает достаточно серьезный уровень интереса к данной тематической области.

2. Региональный анализ поисковых запросов по рассматриваемой тематике позволил выявить неравномерное распределение поисковых запросов по регионам России, что косвенно отражает различную активность регионов России в изучении рассматриваемых нами тем. Естественно, лидируют ведущие научные центры — Москва и Петербург.

3. Обзор российских и зарубежных исследовательских центров

На данном этапе исследования был определен круг исследовательских структур, занимающихся анализом процессов в социальных сетях.

В ходе исследования были проанализированы исследовательские направления деятельности таких российских структур как: Лаборатория интернет-исследований НИУ-ВШЭ-Санкт-Петербург, Тамбовский государственный университет им. Г.Р. Державина, Научно-исследовательская лаборатория социально-гуманитарных исследований и образовательных практик RATIO (Нижегородский государственный педагогический университет), Центр изучения интернета и общества при Российской экономической школе, Научно-исследовательская лаборатория социальных проблем современного общества Центра консалтинга, социологических и маркетинговых исследований «РЕЙТИНГ» (Социологический факультет МГУ), Научно-исследовательская

лаборатория математического моделирования социальных процессов Социологического факультета МГУ, Кафедра философии, социологии и политологии Московского государственного института радиотехники, электроники и автоматики, лаборатория Semantic Future, Институт технологий образования, Научно-исследовательская социологическая лаборатория Курского государственного университета.

Среди зарубежных структур, имеющих схожие исследовательские практики, были отобраны: International Network for Social Network Analysis (INSNA) Department of Sociology, University of South Florida; LINKS Center for Social Network Analysis Gatton College of Business & Economics University of Kentucky; Center for Computational Analysis of Social and Organizational Systems (CASOS); Компания orgnet.com; Simon Fraser University The Vancouver Network Analysis Team; BMBF Innovation Initiative for the New German Länder- Entrepreneurial Regions.

Анализ деятельности исследовательских центров позволил сделать вывод о том, что в настоящее время не существует информационного центра проведения исследований процессов социодинамики на основе веб-платформы, предоставляющей доступ к комплексу автоматизированных программных средств. Вместе с этим масштаб и востребованность проводимых исследований (в том числе в социальных сетях) подтверждает утверждение о том, что на современном этапе эти исследования совершенно невозможно проводить «вручную», не используя автоматизированные инструменты. Среди инструментов, необходимых для исследователей, были выделены следующие: набор исследовательских инструментов в сфере изучения социальных сетей, возможности интерактивного обучения, прохождения лабораторных работ, предоставления информации о себе другим членам профессионального интернет-сообщества, возможности обмена профессиональными материалами с другими членами профессионального интернет-сообщества в рамках виртуальных групп в социальной сети, а также представления своих проектов для общего доступа, для ограниченной совместной работы либо их размещение в закрытом пространстве личного кабинета.

Абсолютное большинство организаций не предоставляют сервиса поддержки среды информационного взаимодействия между своими пользователями. Ни один из центров не поддерживает социальную сеть исследователей.

4. Анализ российских и зарубежных курсов/модулей обучения

В рамках данного исследования был проведен анализ учебных курсов и модулей по социальному моделированию, применению компьютерных технологий в социологии, анализу социальных сетей, предлагаемых российскими и мировыми образовательными центрами. Включение таких курсов в учебные планы программ бакалавриата и магистратуры является признаком востребованности специа-

листов и преподавателей, компетентных в этих вопросах.

При анализе зарубежного опыта были отобраны 11 образовательных курсов, тематика которых лежит в области социального моделирования, применения компьютерных технологий в социологии, анализа социальных сетей.

Большинство курсов по указанной тематике преподаются в университетах США (косвенным показателем этого были высокие показатели количества англоязычных поисковых запросов в данной тематике). Вместе с тем, тематика курсов и их описания позволяют говорить об акценте этих курсов на отдельные предметные области, покрывающие либо широкие теоретические аспекты, либо, наоборот, описывающие узкие направления исследований:

1. Курсы, посвященные статистическим и математическим методам преподаются в Университете Брауна, Йельском университете, Корнелльском университете, Калифорнийском университете в Лос-Анджелесе, Университете Хельсинки.

2. Курсы, посвященные анализу социальных сетей (преимущественно теоретические аспекты): Стэнфордский, Гарвардский, Оксфордский университеты, Австралийский национальный университет.

3. Курсы, посвященные моделированию в социологических исследованиях преподаются в Швейцарской высшей технической школе Цюриха.

Определенные трудности при анализе данных курсов были связаны с тем, что на официальных сайтах не опубликованы подробные программы курсов, а учебные материалы отсутствуют в свободном доступе. Общая тенденция свидетельствует о том, что данные курсы читаются либо на факультетах социологии, либо междисциплинарных учебных центрах. Вышеперечисленные курсы входят в учебный план старших курсов бакалавриата, магистратуры или аспирантуры.

В российской образовательной практике присутствует схожее разделение учебных курсов по тематической направленности:

1. Методы математики и статистики: НИУ Высшая школа экономики, Российский государственный гуманитарный факультет, Южный федеральный университет, Амурский государственный университет

2. Анализ социальных сетей: НИУ Высшая школа экономики,

3. Моделирование социальных процессов: НИУ Высшая школа экономики, Омский, Тюменский, Санкт-Петербургский, Амурский государственные университеты, Российский гуманитарный интернет-университет, Волгоградский центр социальных исследований, Южный федеральный университет

В данном исследовании внимание было уделено именно соответствию содержанию курса тематике связанной с методами проведения исследований в социальных сетях, технологиями контент-анализа, построения моделей социальных процессов и т.д.

Проведенный анализ показал, что изучение исследовательских методик в части моделирования

социальных процессов, изучения взаимодействий в социальных сетях и т.п. должно быть неразрывно связано с практическим освоением материала, выполнением прикладных лабораторных работ, возможностями получения интерактивных консультаций в рамках обучения. На сегодняшний момент рынок образовательных услуг в этой области перенасыщен теоретическими курсами и не обеспечен практикой.

Следует отметить, что выявленные учебные курсы имеют широкую направленность и ориентированы на целый ряд направлений подготовки: социология, экономика, журналистика, прикладная математика и информатика, регионоведение, психология, конфликтология, мировая культура. Вместе с тем теоретическая часть курса в большинстве случаев значительно преобладает над практической либо и вовсе ее замещает.

Таким образом, проведенное исследование позволило зафиксировать интерес к тематике социального моделирования, изучения социальных сетей как в русскоязычном, так и англоязычном сегментах Интернета, востребованность учебных курсов в области социальных сетей и моделирования социальных процессов. Анализ публикуемых научных статей демонстрирует устойчивый интерес исследователей в области гуманитарных наук к применению средств автоматизированного сбора, обработки данных и проведения аналитики.

5. Создание веб-ориентированного центра в области социодинамики и ее приложений

Проведенный комплекс исследований привел к выводам о востребованности автоматизированного веб-центра по проведению исследований процессов, происходящих в социальных сетях.

В Санкт-Петербургском национальном исследовательском университете информационных технологий, механики и оптики (НИУ ИТМО) в 2011 году была начата реализация проекта «Разработка веб-ориентированного производственно-исследовательского центра в области социодинамики и ее приложений» [1]. В рамках проекта был создан веб-центр на базе платформы CLAVIRE (<http://socio.escience.ifmo.ru/>).

Возможности центра позволяют проводить исследования обсуждения различных тем в блогосфере LiveJournal, объединения пользователей в неформальные сообщества, оценивать циркуляцию информационных сообщений между ними. Кроме этого в комплекс встроен учебный модуль, содержащий как теоретические лекционные материалы, так и практические задания и электронные тьюторы.

В ходе реализации проекта и апробации возможностей веб-центра были реализованы несколько исследовательских сценариев [2], в частности:

- построение социогаммы неформального сообщества, анализ скорости и каналов распространения информации;

– анализ и прогноз индексов общественных настроений, мониторинг манипуляций мнениями.

Для реализации данного сценария использовались данные по теме «Законы о пропаганде гомосексуализма и педофилии», принятие которых в регионах России вызвало волну обсуждений, в основном со стороны правозащитников и собственно представителей сексуальных меньшинств, так как, с одной стороны, эти законы практически бесполезны, так как любая сексуальная пропаганда несовершеннолетним (на защиту которых, в основном, направлены законы) и так запрещена законами. С другой стороны, этот закон фактически запрещает проведение гей-парадов и общественных гей-акций, что противоречит международным правовым нормам, в результате чего Европейский суд по правам человека в Страсбурге уже присудил Правительству РФ выплатить штраф в размере 50 тыс. евро за незаконное ограничение прав сексуальных меньшинств по факту отказов в проведении гей-парадов и общественных гей-акций [3]. Для исследования важным фактором является то факт, что данная выборка пользователей может рассматриваться и как сообщество, и как сеть коммуникационного процесса по передаче информации.

На основе ключевых слов по заданной тематике выполнялся краулинг социальной сети «Живой журнал» (ЖЖ), были сформированы сети и совокупности сообщений пользователей, содержащие упомянутые ключевые слова. Сеть составлялась на основе всех найденных сообщений с ключевыми словами и связей подписки между пользователями. Тем не менее, для визуализации и анализа выбирались только актуализированные в данной тематике пользователи сети, относительно недавно писавшие что-то на заданную тему (анализировалось содержание последних 25 записей пользователя).

В состав Центра встроены несколько программных пакетов для анализа сетей и визуализации. На рис. 1 представлен результат построения визуализированной социогаммы сообщества обсуждения темы, визуализированной с помощью программного пакета JUNG.

В ходе анализа полученного неформального сообщества пользователей, интересующихся тематикой «Законы о пропаганде гомосексуализма» были получены следующие результаты:

1) В сообществе относительно мало одиночек. Одиночные кружки/точки обозначают пользователей, не связанных подпиской с другими членами сообщества.

2) Явно проявляются процессы сетевой самоорганизации, охватывающие достаточно большое количество членов этого неформального сообщества.

3) Исследование позволило выделить наиболее влиятельных членов сообщества, лидеров мнений – по количеству сообщений, которые были «подхвачены» подписанными на этих лидеров мнений членами сообщества.

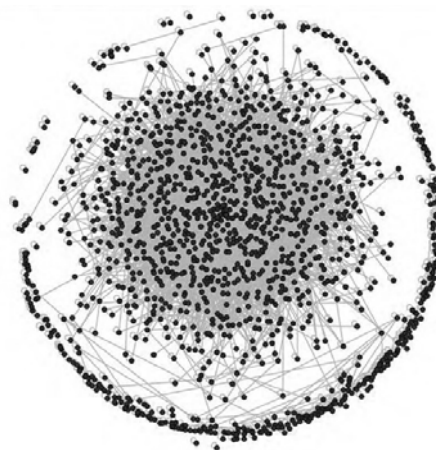


Рис. 1. Социогамма неформального сообщества пользователей ЖЖ, обсуждающих тематику «Законы о пропаганде гомосексуализма и педофилии» (получено с использованием пакета JUNG)

4) С помощью программных средств web-центра были рассчитаны характеристики частоты появления сообщений по тематике «Законы о пропаганде гомосексуализма».

В настоящее время в научной литературе и на конференциях активно обсуждается факт большого влияния социальных сетей на распространение информации, в том числе – социально-значимой и потенциально опасной. В то же время не вполне понятен масштаб проблемы (оценочная численность вовлеченных в распространение такого рода информации) и уровень взаимосвязанности (либо разрозненности) пользователей социальных сетей, участвующих в такого рода взаимодействиях. Социальные сети становятся все более массовым и быстрым каналом распространения информации, когда за счет перепоста сообщений оно может распространяться очень быстро на очень большие аудитории.

5. Заключение

Современная исследовательская проблема заключается в необходимости объяснения процессов трансформации «реального» социального пространства в связи с появлением новых «виртуальных» институциональных социальных явлений и социальных групп. Для решения данной проблемы необходимо развитие инструментов социометрии прикладного характера.

Обозначенные научные проблемы с октября 2012 г. решаются исследовательской группой в рамках проекта «Исследование специфики распространения социально-значимой информации в социальных сетях на примере Живого Журнала с помощью автоматизированных программных средств» (грант Минобрнауки РФ № 2012.-1.2.2-12-000-3003-056).

Реализация данного исследовательского проекта позволит решить следующие исследовательские задачи:

– выявление с помощью автоматизированных программных средств и визуализация в виде ориентированных графов социогамм виртуальных

каналов социально-значимой и потенциально опасной информации;

- анализ скорости распространения информации в социальных сетях;
- выявление наиболее частых источников подобной социально-значимой и потенциально опасной информации, составление списка таких источников;
- выявление динамики индексов актуализированности в общественном сознании социальных проблем как на примере частоты упоминания соответствующих ключевых слов в записях Живого журнала, так и сравнение этих индексов с динамикой частоты запросов тех же ключевых слов в поисковых системах (на примере статистики ключевых слов Яндекс);
- проверка гипотезы о корреляции данных динамики статистики соответствующих ключевых слов в поисковых запросах Яндекса и динамики статистики ключевых слов записей Живого Журнала.

Литература

- [1] Иванов С. В., Болгова Е. В., Каширин В. В., Якушев А. В., Чугунов А. В., Бухановский А. В. Web-ориентированный производственно-исследовательский центр „Социодинамика“ // Известия высших учебных заведений. Приборостроение. 2011. №10. С. 65- 72.
- [2] Бершадская Л.А., Биккулов А.С., Болгова Е.В., Чугунов А.В. Социальные сети и социометрические исследования: теоретические основания и практика использования автоматизированного инструментария изучения виртуальных сообществ // Информационные ресурсы России. - 2012. № 4. С. 19-24.
- [3] Петербургский закон о пропаганде гомосексуализма обжалован в Страсбургском суде // Деловой Петербург. 13.06.2012. [Электронный ресурс]. URL: http://www.dp.ru/a/2012/06/13/Peterburgskij_zakon_o_pro/ (дата обращения 20.07.2012).

The need in automated tools on humanities research market

L. Bershadskaya, A. Bikkulov

Marketing research results in the field of automated tools in humanities are given in the article. The authors explain the urgency and relevance of web-based scientific center in the field of sociodynamic processes research.

* Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 - 2013 годы» по теме: «Исследование специфики распространения социально-значимой информации в социальных сетях на примере Живого Журнала с помощью автоматизированных программных средств».