



LABORATORY FOR INTERNET STUDIES



www.linis.ru

PEOPLE

**Olessia
Koltsova,
Head of
laboratory**



**Sergey
Nikolenko,
Mathematician,
Senior
researcher**



**Sergei Koltcov,
Physicist,
technical
director**

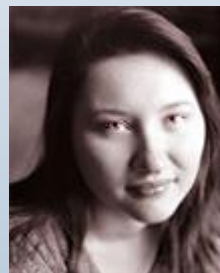


**Peter Meylakhs,
Sociologist,
Senior
researcher**



**Anastasia
Shimorina,
Computer linguist,
Junior researcher**

**Svetlana
Bodrunova,
Media scholar,
Invited
researcher**



**Andrey Scherbak,
Political scientist,
Invited researcher**



**Yury Rykov,
PhD student in
sociology,
intern**



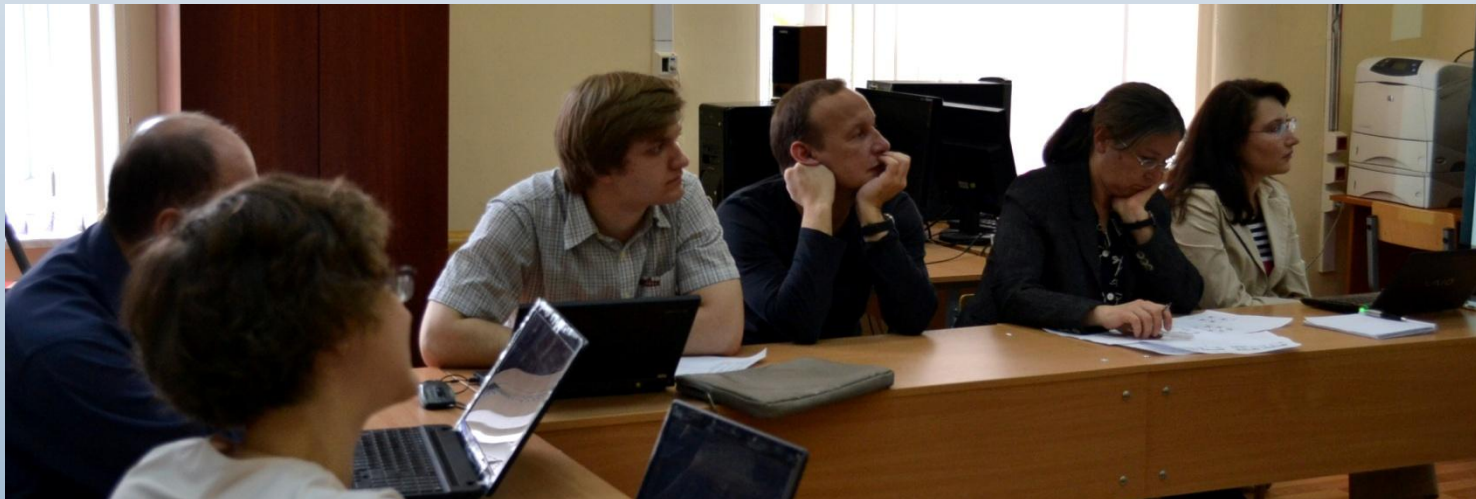
**Tatiana Yefimova,
PhD student in
political science,
intern**



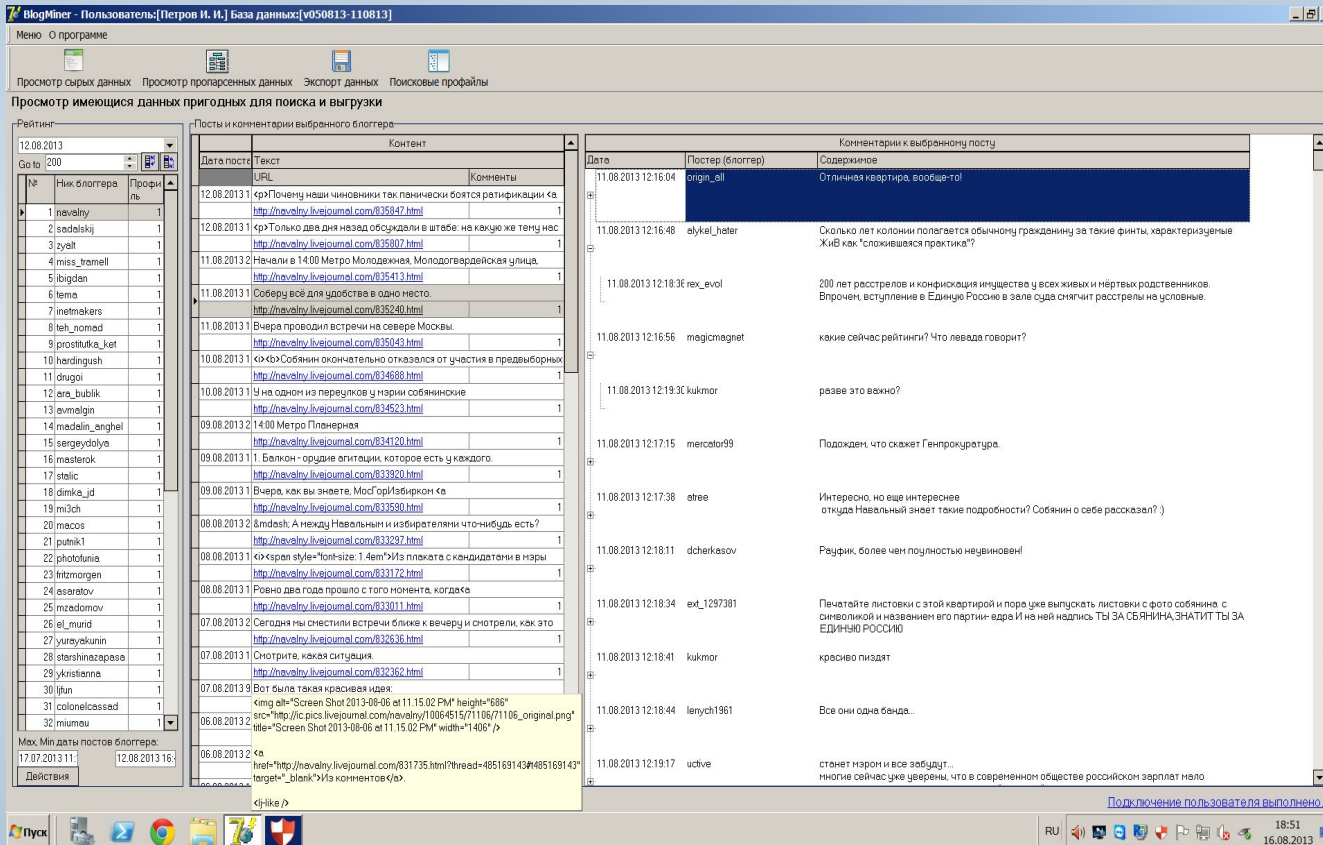
**Daria
Yudenkova
Administrative
director**

LAB'S THREE CORNERSTONES

- Big Internet data
- Diverse methods & tools
- Social science research questions and goals



DATA1: LIVEJOURNAL



BlogMiner - Пользователь:[Петров И. И.] база данных:[v050813-110813]

Меню О программе

Просмотр сырых данных Просмотр пропаренных данных Экспорт данных Поисквые профайлы


Просмотр имеющихся данных пригодных для поиска и выгрузки

Рейтинг: 12.08.2013

Go to: 200

| № | Ник блоггера | Пройриль |
|----|-----------------|----------|
| 1 | navalny | 1 |
| 2 | sadelskij | 1 |
| 3 | zyelt | 1 |
| 4 | miss_tromell | 1 |
| 5 | ibigidan | 1 |
| 6 | tema | 1 |
| 7 | inetmakers | 1 |
| 8 | leh_nomad | 1 |
| 9 | prostotka_ket | 1 |
| 10 | herdingush | 1 |
| 11 | drugoi | 1 |
| 12 | era_bublik | 1 |
| 13 | avmalgin | 1 |
| 14 | madalin_angel | 1 |
| 15 | sergeyvolya | 1 |
| 16 | maslerok | 1 |
| 17 | stalic | 1 |
| 18 | dinka_id | 1 |
| 19 | mi3ch | 1 |
| 20 | macos | 1 |
| 21 | putnik1 | 1 |
| 22 | photofonia | 1 |
| 23 | hitzmorgen | 1 |
| 24 | eserstov | 1 |
| 25 | mzedomov | 1 |
| 26 | el_muid | 1 |
| 27 | yuryyukunin | 1 |
| 28 | starshinaterepa | 1 |
| 29 | ykrishtanna | 1 |
| 30 | lfln | 1 |
| 31 | colonelcassad | 1 |
| 32 | miumau | 1 |

Max. Мин даты постов блоггера:
17.07.2013 11: 12.08.2013 16:
Действия

| Дата поста | Текст | Комменты |
|--------------|---|---|
| 12.08.2013 1 | кр?Почему наши чиновники так панически боятся ратификации € | http://navalny.livejournal.com/835947.html |
| 12.08.2013 1 | кр?Только два дня назад обсуждали в штебе на какую же тему нас | http://navalny.livejournal.com/835907.html |
| 11.08.2013 2 | Начали в 14:00 Метро Молодежная. Молодогвардейская улица. | http://navalny.livejournal.com/835413.html |
| 11.08.2013 1 | Соберу всё для удобства в одно место. | http://navalny.livejournal.com/835240.html |
| 11.08.2013 1 | Вчера проводил встречи на севере Москвы. | http://navalny.livejournal.com/835043.html |
| 10.08.2013 1 | «»Собянин окончательно отказался от участия в предвыборных | http://navalny.livejournal.com/834688.html |
| 10.08.2013 1 | Ч на одном из переулков у мэрии собянинские | http://navalny.livejournal.com/834523.html |
| 09.08.2013 2 | 14:00 Метро Плевнерная | http://navalny.livejournal.com/834120.html |
| 09.08.2013 1 | 1. Балкон - орудие агитации, которое есть у каждого. | http://navalny.livejournal.com/833920.html |
| 09.08.2013 1 | Вчера, как вы знаете, МосГор избирком € | http://navalny.livejournal.com/833590.html |
| 08.08.2013 2 | — А между Навальным и избирателями что-нибудь есть? | http://navalny.livejournal.com/833297.html |
| 08.08.2013 1 | «»Круп style="font-size: 1.4em">Из плаката с кандидатами в мэри | http://navalny.livejournal.com/833172.html |
| 08.08.2013 1 | Ровно два года прошло с того момента, когда € | http://navalny.livejournal.com/833011.html |
| 07.08.2013 2 | Сегодня ны снесли вострой ближе к вечеру и снятели, как это | http://navalny.livejournal.com/832636.html |
| 07.08.2013 1 | Снотрите, какая ситуация | http://navalny.livejournal.com/832362.html |
| 07.08.2013 9 | Вот была такая красивая идея |  |
| 06.08.2013 2 | «» | http://navalny.livejournal.com/831735.html?head=465169143#465169143 |

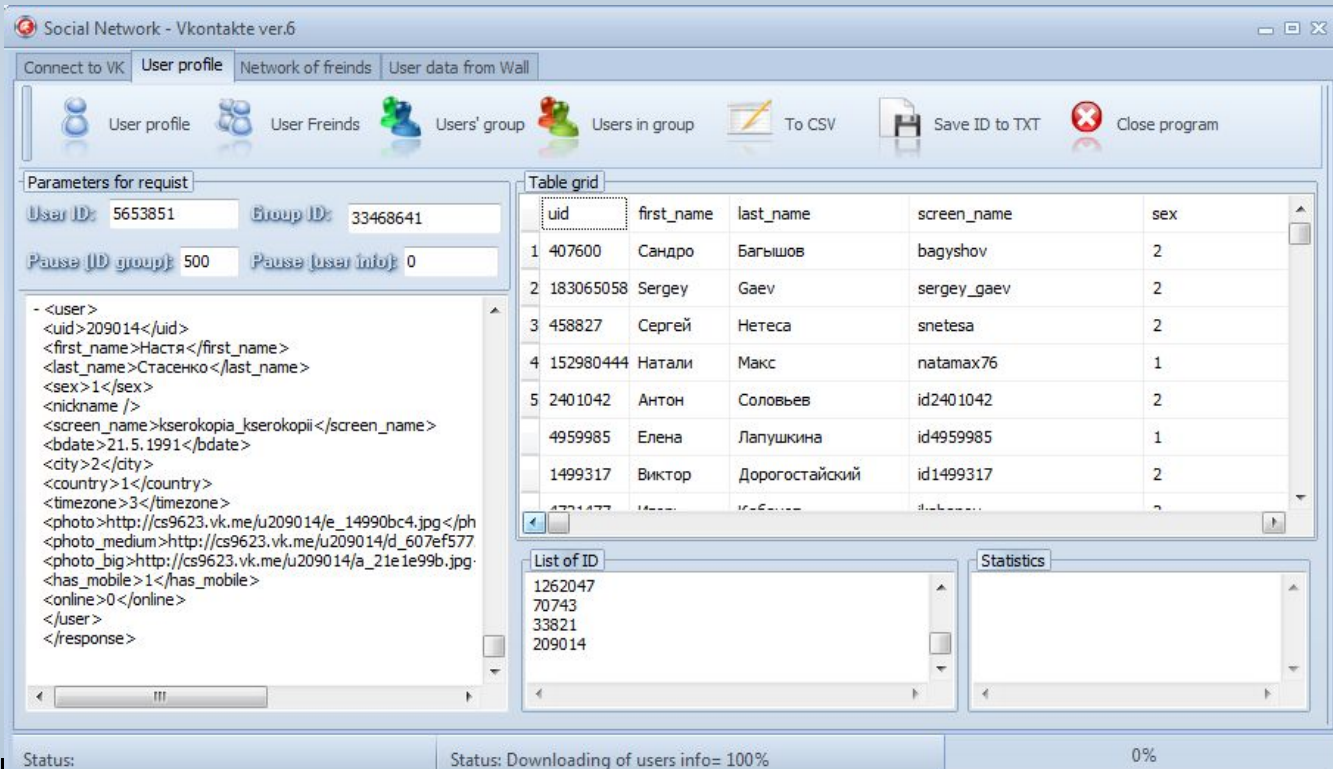
| Дата | Постер (блоггер) | Содержимое |
|---------------------|------------------|---|
| 11.08.2013 12:16:04 | orig_nal | Отличная квартира, вообще-то! |
| 11.08.2013 12:16:48 | ajkal_hater | Сколько лет колонии полагается обычному гражданину за такие финты, характеризующие Жиб как "сплохишевая практика"? |
| 11.08.2013 12:18:36 | tek_evil | 200 лет расстрелов и конфискации имущества у всех живых и мёртвых родственников. Впрочем, вступление в Единую Россию в зале суда снимчит расстрелы на услове. |
| 11.08.2013 12:16:56 | mejsmagnet | какие сейчас рейтинги? Что левада говорит? |
| 11.08.2013 12:19:30 | kukmor | разве это вино? |
| 11.08.2013 12:17:15 | mercato99 | Подожден, что скажет Генпрокуратура. |
| 11.08.2013 12:17:38 | afree | Интересно, но еще интереснее откуда Навальный знает такие подробности? Собянин о себе рассказал?) |
| 11.08.2013 12:18:11 | dchekasov | Реуцик, более чем полностью неувинен! |
| 11.08.2013 12:18:34 | ex_1297381 | Печатайте листовки с этой квартирой и поре уже выпускайте листовки с фото собянина, с символикой и названием его парти-вдра И на ней надпись ТЫ ЗА СЕЯНИНА ЗНАТИТ ТЫ ЗА ЕДИНУЮ РОССИЮ |
| 11.08.2013 12:18:41 | kukmor | красиво пиздят |
| 11.08.2013 12:18:44 | lenych1961 | Все они одна банд... |
| 11.08.2013 12:19:17 | active | станет мэром и все забудут... многие сейчас уже уверены, что в современном обществе российский зарплат мало |

Подключение пользователя выполнено.

RU 18:51 16.08.2013

Koltran BLOGMINER mines posts, comments and user metadata

DATA2: VKontakte SOCIAL NETWORK



Connect to VK User profile Network of freinds User data from Wall

User profile User Freinds Users' group Users in group To CSV Save ID to TXT Close program

Parameters for requist

User ID: 5653851 Group ID: 33468641

Pause (ID group): 500 Pause (user info): 0

```

- <user >
<uid>209014</uid>
<first_name>Настя</first_name>
<last_name>Стасенко</last_name>
<sex>1</sex>
<nickname />
<screen_name>kserokopia_kserokopii</screen_name>
<bdate>21.5.1991</bdate>
<city>2</city>
<country>1</country>
<timezone>3</timezone>
<photo>http://cs9623.vk.me/u209014/e_14990bc4.jpg</ph
<photo_medium>http://cs9623.vk.me/u209014/d_607ef577
<photo_big>http://cs9623.vk.me/u209014/a_21e1e99b.jpg
<has_mobile>1</has_mobile>
<online>0</online>
</user >
</response >

```

| uid | first_name | last_name | screen_name | sex |
|-------------|------------|----------------|-------------|-----|
| 1 407600 | Сандро | Багышов | bagyshov | 2 |
| 2 183065058 | Sergey | Gaev | sergey_gaev | 2 |
| 3 458827 | Сергей | Нетеса | snetesa | 2 |
| 4 152980444 | Натали | Макс | natamax76 | 1 |
| 5 2401042 | Антон | Соловьев | id2401042 | 2 |
| 4959985 | Елена | Лапушкина | id4959985 | 1 |
| 1499317 | Виктор | Дорогостайский | id1499317 | 2 |

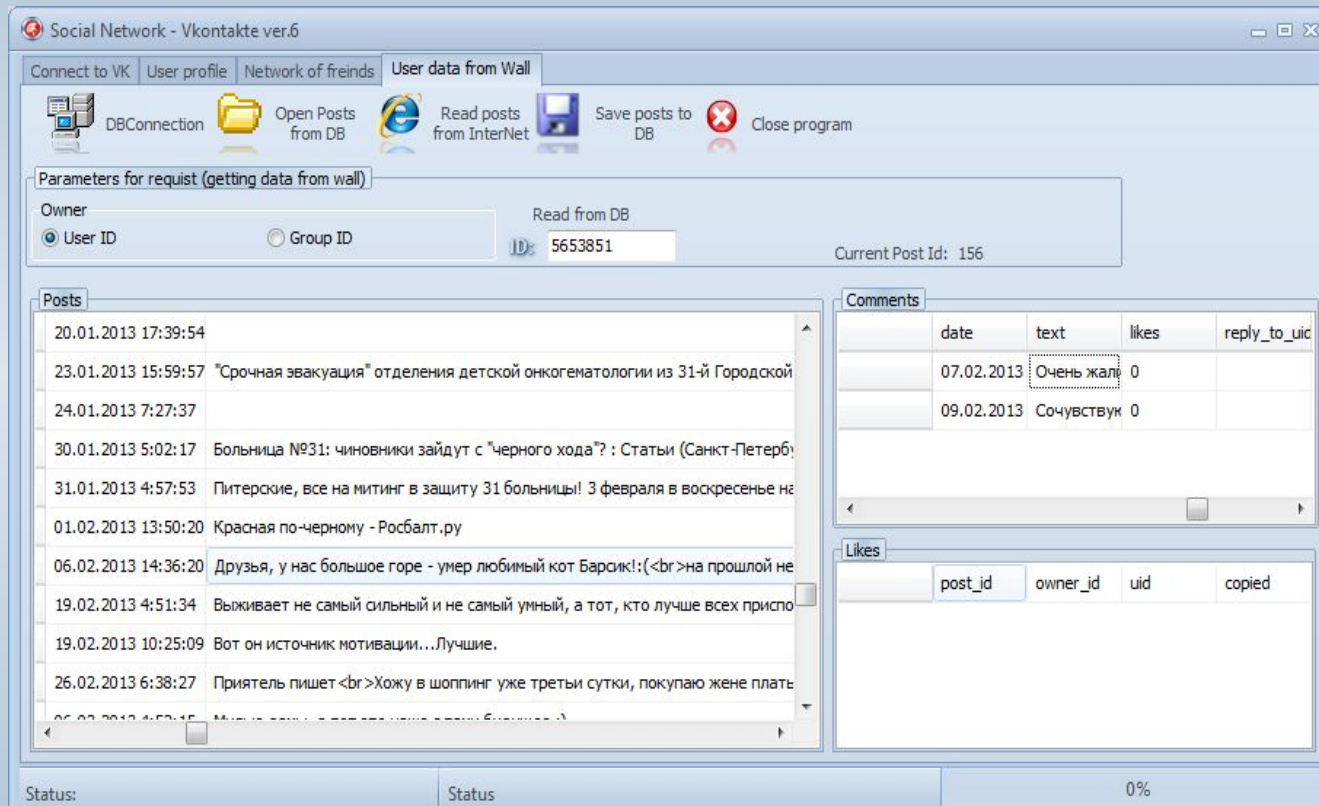
List of ID
1262047
70743
33821
209014

Statistics

Status: Status: Downloading of users info= 100% 0%

LINIS vkminer mines friendship networks, friendship networks of VK communities and....

DATA2: VKontakte SOCIAL NETWORK



...and message boards of VK communities (posts, comments, likes, user metadata).

TOOLS & METHODS 1

Text mining: topic modeling with LINIS TopicMiner

TopicMiner ver. 41

Stemming | Gibbs LDA sampling | Kullback Leibler Distance

Parameters of simulation

Alpha: 0,5
Beta: 0,1
Save step: 50

Number of documents: Label1

Perplexity

Status: Calculation completed

Calculation method

Words with high probability

| | 1 | 2 | 3 | 4 | 5 | 6 |
|----|-------------------|-------------------|----------------------|--------------------|---------------------|--------------------|
| 1 | Федерация: 0,0147 | верба: 0,007033 | мрамор: 0,015868 | mdash: 0,016396 | танк: 0,018219 | японский: 0,015... |
| 2 | российский: 0,012 | ад: 0,005830 | соотечественник: 0,0 | мост: 0,008951 | т: 0,006409 | фонд: 0,010658 |
| 3 | трудоной: 0,00778 | хусаиновна: 0,00 | ваниль: 0,007334 | остров: 0,008124 | мм: 0,006021 | япония: 0,00916; |
| 4 | к: 0,007614 | женщина: 0,0044 | стручок: 0,004287 | тоннель: 0,006304 | вооружение: 0,00544 | только: 0,008497 |
| 5 | работодатель: 0,0 | ангелин: 0,00422 | месторождение: 0,00 | остер: 0,005642 | машин: 0,005247 | вот: 0,008331 |
| 6 | район: 0,006578 | террорист: 0,0036 | гражданство: 0,00406 | поезд: 0,005477 | пушка: 0,004666 | посольство: 0,00 |
| 7 | статья: 0,006233 | ветка: 0,003226 | камень: 0,004084 | строительство: 0,0 | название: 0,004472 | деньга: 0,006834 |
| 8 | работник: 0,00606 | земля: 0,003025 | разный: 0,003271 | железнодорожный | башня: 0,004279 | пономарев: 0,00 |
| 9 | расход: 0,005715 | дерево: 0,002825 | нация: 0,003068 | шоколад: 0,002995 | польский: 0,003504 | нужный: 0,00650 |
| 10 | стоимость: 0,0057 | почка: 0,002625 | мраморный: 0,002865 | км: 0,002995 | армия: 0,003504 | американский: 0, |
| 11 | север: 0,005369 | освящать: 0,0022 | гражданин: 0,002865 | бисквит: 0,002829 | украинский: 0,00331 | британский: 0,00 |
| 12 | отпуск: 0,005369 | душ: 0,002024 | РВП: 0,002662 | стоимость: 0,0028 | снаряд: 0,003117 | правда: 0,00550 |
| 13 | компенсация: 0,00 | веточка: 0,00182 | россия: 0,002662 | шоколадный: 0,002 | комплекс: 0,003117 | японец: 0,00517 |
| 14 | конституционный: | век: 0,001823 | документ: 0,002458 | тут: 0,002664 | орудие: 0,002924 | пекин: 0,005005 |
| 15 | крайний: 0,005024 | вход: 0,001823 | квота: 0,002255 | линия: 0,002498 | платформа: 0,00234 | восток: 0,005005 |
| 16 | договор: 0,004852 | праздник: 0,0018 | полчение: 0,002255 | участок: 0,002333 | тавар: 0,002149 | весел: 0,004835 |

Export to Excel... Number of words for export: 100

TOOLS & METHODS 2

Text mining: sentiment analysis

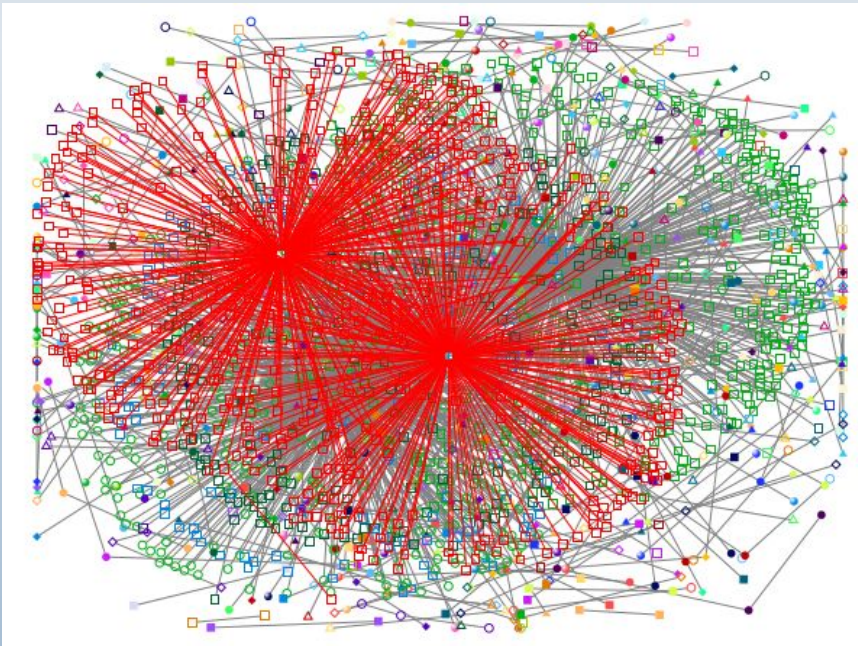
Russian language
tools still need to be
developed

Sentiment lexicon will
be constructed and
tested with
SentiStrength (Thelwall)



TOOLS & METHODS 3

Network analysis



Many tools exist, but most do not process data of our scale

Community detection with Louvain source code

RESEARCH QUESTIONS AND GOALS 1

- Online public agenda: topical structure of the Russian blogosphere and influence of national elections 2011-2012 on it (finished project)



RESEARCH QUESTIONS AND GOALS 3



НЕАНДЕРТАЛЬЦЫ
теперь у тебя во дворе

demotivation.ru

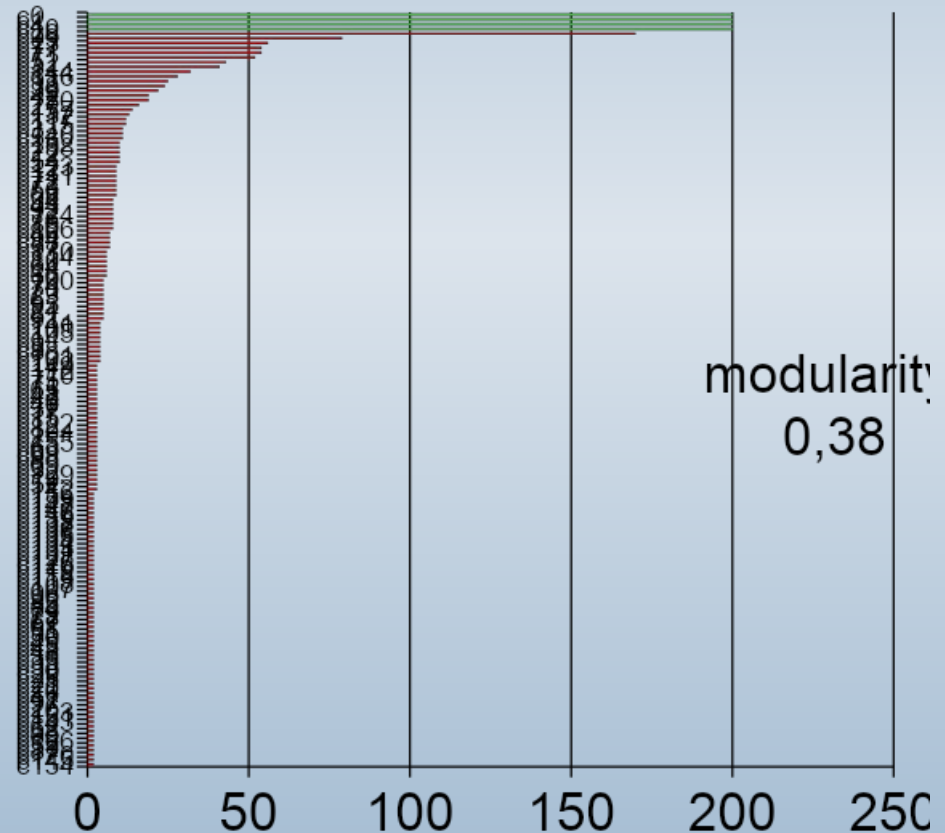
“Neanderthals are now in your backyard” (a picture from a blog post concerning Muslims slaughtering sheep for Kurban-Bayram in Moscow)

Online public opinion:
revealing topics that
arouse social tension
through sentiment
analysis

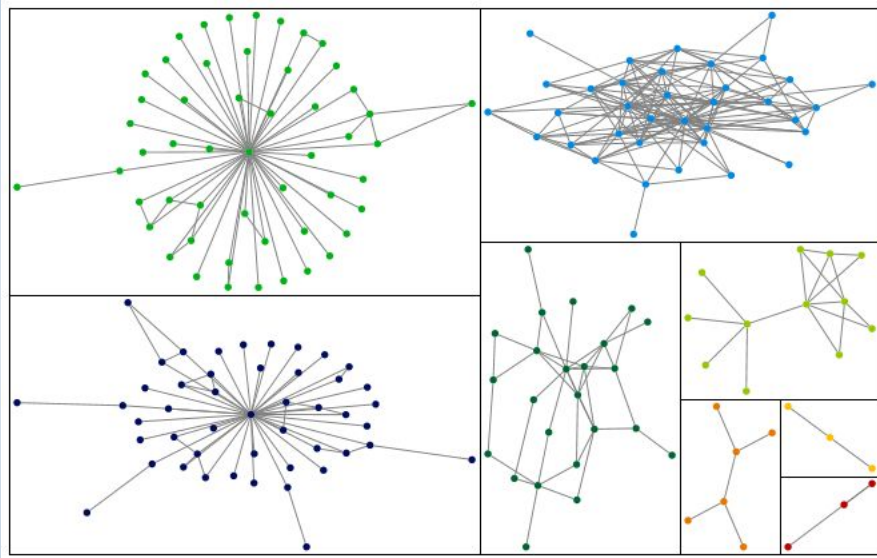
(pilot project finished;
main project planned)

RESEARCH QUESTIONS AND GOALS 3

Comment-based
communities,
their topical
coherence and
authorship
homogeneity
(finished project)

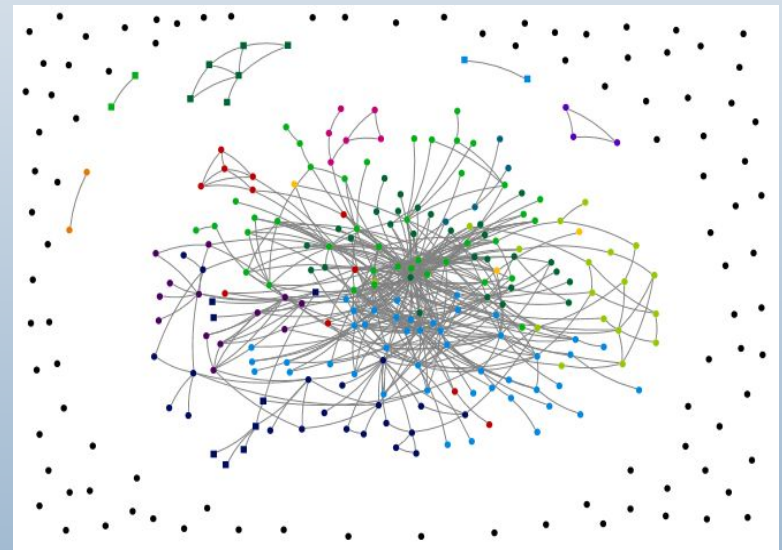


RESEARCH QUESTIONS AND GOALS 4



Network structure
and leadership in VK
communities of
social movements

Influentials and
suscep-tibles in
destructive com-munities:
perspectives for Internet



THANK YOU!



- Главная
- О нас
- События
- Публикации
- Исследования
- Материалы
- Ссылки
- Контакты

ENGLISH



Новости



Чем дышит блогосфера? Семинар ЛИНИс в Москве

11:54 27.04.2012

24 апреля команда ЛИНИс провела в НИУ-ВШЭ (Москва) семинар на тему "Чем дышит блогосфера? Методы анализа больших массивов Интернет-данных для социологических задач". Выступление прошло в рамках академического семинара по социологической теории и методологии кафедры анализа социальных институтов (рук. Инна Девятко), который на этот раз проходил совместно с Лабораторией экономико-социологических исследований (рук. В.Радаев).

[Подробнее...](#)

ЛИНИс – участник Балтийского партнерства по новым медиа



Анонсы

11.05.2012 (Пятница)

Презентация Лаборатории Интернет-Исследований

11 мая в 17-00 состоится презентация новой исследовательской площадки Национального исследовательского университета «Высшая школа экономики» (Санкт-Петербург) - Лаборатории интернет-исследований (ЛИНИс).

[Подробнее...](#)

26.09.2012 (Среда)

Новые СМИ: меняющийся медийный ландшафт

27-28 сентября 2012 года лаборатория интернет-исследований совместно с зарубежными партнерами проводит конференцию "New media: changing media landscapes". Конференция