

Исследование порталов электронного участия в России на примере извлечения данных с открытых площадок

Б.А. Низомутдинов, А.С. Тропников

Университет ИТМО

boris-wels@yandex.ru, flairwow@gmail.com

Аннотация

Электронное участие становится все более важным фактором развития успешных отношений между государством и обществом. Существуют различные способы электронного участия. Один из них - обращение к власти через интернет-приемную. В статье представлены результаты пилотного исследования, предполагающего контент-анализ обращений граждан. Авторы собрали список интернет-приемных, на которых находятся в открытом доступе тексты обращений — это обращение граждан с ответами. Из общего списка было отобрано 5 площадок из разных регионов с максимальным количеством обращений, на основе которых проведен контент-анализ опубликованных заявок. Результаты представленного исследования могут быть полезны для разработки методологии и критериев оценки электронного участия в России.

Ключевые слова: электронное участие, интернет приемная, обращения, персональные данные

Библиографическая ссылка: Низомутдинов Б.А., Тропников А.С. Исследование порталов электронного участия в России на примере извлечения данных с открытых площадок // Государство и граждане в электронной среде. Выпуск 3 (Труды XXII Международной объединенной научной конференции «Интернет и современное общество», IMS-2019, Санкт-Петербург, 19–22 июня 2019 г. Сборник научных трудов). — СПб: Университет ИТМО, 2019. С. 155 – 163. DOI: 10.17586/2541-979X-2019-3-155-163

1. Введение

В мировой практике электронное участие граждан в политической жизни является важным элементом в развитии современного демократического общества. В научной литературе имеются данные о тех факторах, которые оказывают положительное влияние на адаптацию и использование новых инструментов электронного участия.

Электронное участие — это интернет-механизмы учета мнения населения при принятии государственных решений. К ним относятся электронные приемные, порталы для онлайн-голосований и общественных инициатив, специализированные порталы, которые предназначены для оперативного решения городских проблем — все то, что позволяет гражданам общаться с чиновниками, контролировать их деятельность и влиять на публичную политику.

Коллектив авторов [1] рассматривает вопросы выработки критериев для оценки социальной эффективности порталов электронного участия в России. Исследователи [2] разработали методику оценки электронного участия граждан в публичной политике, так, для тестирования технологии было проведено исследование электронных приемных российских регионов, по результатам которого был сделан вывод, что они предоставляют относительно ограниченные возможности для контроля за деятельностью власти, а некоторые регионы включены в электронное участие чисто формально. Согласно новому методу, эффективность электронного участия определяется не техническими

возможностями, а институциональным дизайном интернет-сайтов. Для анализа электронных петиций в Университете ИТМО была разработана автоматизированная система аналитики порталов электронных петиций (<http://analytics.egov.ifmo.ru/>). Система позволяет в режиме реального времени отслеживать появление новых петиций по различным тематикам, определять ход голосования по ним, а также строить прогноз относительно того, преодолит ли она необходимый порог голосов для перехода на обсуждение в экспертную группу. Группы исследователей изучают вовлеченность граждан в процессы электронного участия [3], анализируют влияние сообществ в социальных сетях на публикацию электронных петиций и их распространение. По данным последних исследований [4], один из факторов, способствующих развитию инструментов электронного участия являются социальные сети, которые служат катализатором привлечения новых пользователей к площадкам электронного участия. Авторы приводят результаты сравнительного исследования сообществ в социальной сети ВКонтакте, посвященных государственным и негосударственным порталам электронных петиций. Внимание российских исследователей уделяется также анализу публичных дискуссий в интернете на основе концепции делиберативной демократии [5], так, используемые подходы к исследованиям публичного дискурса могут быть адаптированы и к другим исследованиям дискурса в интернете вокруг значимых политических событий.

Однако, чаще всего, данные работы направлены на изучение содержаний петиций, обсуждений и голосований, но малоизученным остается контекст содержания обращений граждан в органы власти через интернет-приемные, а ведь именно изучение текстов таких обращений представляет отдельный интерес и может помочь в выявлении социально значимых проблем.

Интернет-приемная исполнительных органов государственной власти — это сервис по отправке обращений граждан должностным лицам. Данный сервис предоставляет возможность отправить обращение в форме электронного документа в администрацию, комитеты, управления, инспекции, и другие службы, получить информацию о ходе рассмотрения поданных ранее обращений. Как правило, обращения рассматриваются в соответствии с Федеральным законом от 2 мая 2006 года № 59 «О порядке рассмотрения обращений граждан Российской Федерации» [6]. Данным Федеральным законом регулируются правоотношения, связанные с реализацией гражданином Российской Федерации закрепленного за ним Конституцией Российской Федерации права на обращение в государственные органы и органы местного самоуправления, а также устанавливается порядок рассмотрения обращений граждан государственными органами, органами местного самоуправления и должностными лицами. Установленный настоящим законом порядок рассмотрения обращений граждан распространяется на все обращения граждан, за исключением обращений, которые подлежат рассмотрению в порядке, установленном федеральными конституционными законами и иными федеральными законами.

Обращения граждан направляются в соответствующие государственные органы и организации, осуществляющие публично значимые функции, в компетенцию которых входит решение поставленных вопросов. В случае необходимости подтверждения своих доводов гражданин вправе приложить к обращению необходимые документы и материалы в электронной форме.

Информация, текст обращения, направленных в электронном виде, и ответ на них хранится и обрабатывается с соблюдением требований российского законодательства о персональных данных. Ответ на обращение, либо уведомление о его переадресации направляется автору соответствующим должностным лицом или органом исполнительной власти, в зависимости от содержания вопроса, изложенного в обращении, в форме электронного документа либо по почтовому адресу, если автором указаны фамилия, имя, отчество (при наличии), почтовый адрес или адрес электронной почты (e-mail). На обращение, в котором не указаны фамилия заявителя или электронный адрес, по которому должен быть направлен ответ, ответ не дается.

Как правило, текст обращения и ответ на него не публикуется в открытом виде в интернет-приемной, иначе это можно считать разглашением сведений, содержащихся в обращении гражданина. Такая информация предназначена только для органов местного самоуправления или должностных лиц, в компетенцию которых входит решение поставленных в обращении вопросов. К примеру, часто через интернет-приемные задают вопросы о положенных субсидиях, компенсациях и льготах, что можно отнести к личной информации.

Учитывая данный момент, большая часть данных из интернет-приемных закрыта для исследователей, с точки зрения контент-анализа текста. Так, в своей разработке методики оценки электронного участия граждан в публичной политике, политологи [2] провели исследование электронных приемных российских регионов, согласно которому, только 20% изученных приемных имеют опубликованные обращения и ответы. 80% - не публикуют обращения граждан в открытом доступе. Однако, авторы данной работы рассматривали только сайты региональных правительств, полный список интернет-приемных региональных, областных и муниципальных органов власти на данный момент отсутствует.

В интернет-приемных, где нет текстов обращений и ответов, исследователи могут получить лишь статистическую информацию о работе с письменными и устными обращениями граждан, узнать распределение количества вопросов по тематическим разделам, содержащихся в обращениях.

Все обращения граждан, поступающие в интернет-приемную субъекта, как правило классифицируются в соответствии с типовым общероссийским тематическим классификатором обращений граждан, организаций и общественных объединений с отнесением их к соответствующей теме, тематике и тематическому разделу на основании аннотаций обращений. Подобные отчеты публикуются на страницах интернет-приемных.

Согласно Типовому общероссийскому тематическому классификатору обращений граждан, организаций и общественных объединений [7], все обращения граждан можно разделить на разделы:

- государство, общество и политика;
- социальная сфера;
- экономика;
- оборона, безопасность, законность;
- жилищно-коммунальная сфера.

Однако, в ходе анализа, было выявлено, что в сети имеется некоторое количество интернет-приемных органов власти, где в открытом доступе можно найти текст обращений граждан. Таким образом, можно собрать такие обращения, с целью дальнейшего контент-анализа сообщений, кластеризации по темам и др.

Для этого, исследователи проанализировали интернет-ресурсы региональных органов власти и органов местного самоуправления на предмет наличия полных текстов обращений в интернет-приемных.

2. Поиск открытых данных в интернет приемных

Задачей нашего исследования является проработка методов сбора и контент-анализа полнотекстовых обращений в интернет-приемных для выявления наиболее актуальных проблем, которые затрагиваются гражданами при электронном участии. В нашем исследовании предпринята попытка проанализировать возможность анализа полных текстов обращений и ответов на них, используя данные, размещенные в интернет-приемных в открытом доступе. В ходе анализа был проведен поиск и отбор площадок, которые публикуют обращения, организован сбор обращений, предпринята попытка автоматизированного сбора, а также проведен контент анализ обращений и ответов.

Алгоритм анализа включал следующие этапы:

- поиск и отбор площадок с использованием языка запросов Яндекс;

- сбор информации об обращениях;
- сбор обращений;
- статистическая обработка обращений.

Также, в ходе исследования было проведено сравнение площадок по публикационной активности автором. Для анализа площадок использовались следующие данные:

- общее количество обращений;
- дата первой и последней публикаций (постов), общее количество постов.

На следующих этапах исследования планируется проанализировать связи между обращениями, для поиска наиболее активных граждан, а также с целью оценки степени вовлеченности граждан.

В первой части исследования был проведен поиск площадок, на которых есть обращения в открытом доступе, обычным запросом через поисковую систему Яндекс. В поисковой выдаче было выявлено несколько приемных, где опубликованы полные тексты. Все они подверглись анализу для выявления общих признаков, с целью формирования верного поискового запроса для массового поиска. Для точного поиска использовался язык запросов Яндекс. Интернет-приемные без текстов обращений пропускались.

Язык запросов позволяет точнее сформулировать поисковый запрос. Синтаксис языка запросов включает логические и расширенные операторы. Правила запросов Яндекс имеют следующие особенности: список операторов и морфологию, сложные запросы, описание порядка слов, исключение и ограничение расстояния между словами.

Так, поисковые фразы «интернет приемная» расширялись дополнениями, к примеру «добрый день», так как запрос граждан начинается именно с обращения. Такой запрос подразумевает, что мы ищем интернет приемную, и на этой странице уже есть обращения граждан. Согласно нашим наблюдениям, запрос граждан начинается чаще всего со стандартных речевых оборотов, наподобие фразы «добрый день», а в шапке обращения почти всегда есть слово «интернет приемная», тем самым мы можем искать площадки, где открыто публикуют обращения граждан.

В итоге, в ходе работы с поисковой системой, было отобрано 40 интернет приемных различного уровня, в которых в открытом доступе есть полные тексты обращений с данными авторов, все они были классифицированы по регионам, так, самое большое количество площадок было найдено в Челябинской области - 5 штук (см. табл. 1).

Таблица 1. Пример интернет приемных с полными текстами обращений

№	Название площадки	Ссылка
1	Онлайн приёмная администрации Советского района города Челябинска	http://sovadm74.ru/online-priemnaya
2	Официальный сайт Уйского муниципального района	http://xn----8sbwdbcc3airgw.xn--plai/reception/page/1.php
3	Официальный сайт Администрации Курчатовского района	http://kurchatov74.ru/FAQ?FilterCategory=&PageSize=100
4	Администрация Тракторозаводского района города Челябинска	http://admintzr.ru/internet-priemnaya/
5	Официальный сайт администрации Ленинского района	http://lenadmin74.ru/gbook.php?&edge=60

Стоит отдельно отметить, что из итоговой базы интернет-приемных были исключены площадки по решению городских проблем, а также площадки по подаче петиций.

3. Анализ обращений в интернет-приемные

После отбора площадок они были объединены в одну базу, база содержала информацию по количеству обращений на каждой площадке, дату первого и последнего обращения.

Для пилотного исследования было отобрано 5 площадок из разных регионов с максимальным количеством обращений. Список площадок, с которых происходил сбор обращений, представлен в таблице 2.

Таблица 2. Список площадок для анализа обращений

№	Ссылка
1	https://vologda-oblast.ru/onlayn_priyemnaya_gubernatora/
2	http://sovadm74.ru/online-priemnaya
3	http://gradkostroma.ru/feedback/index.aspx
4	http://lenadmin74.ru/gbook.php?&edge=4530
5	http://www.priozersk.lenobl.ru/priemnaya?start=450

Далее, с каждой интернет-приемной из табл. 2 были выгружены обращения, опубликованные за последние 3 месяца. Извлечение информации производилось оператором, который выделял каждое обращение, копировал его и добавлял в таблицу Excel, затем копировал и добавлял ответ.

Следующим шагом стала загрузка корпуса текстов в систему Advego, для семантического анализа. Семантический анализ — этап в последовательности действий алгоритма автоматического понимания текстов, заключающийся в выделении семантических отношений, формировании семантического представления текстов. Семантический анализ текста оценивает количество слов или фраз, которые определяют смысл текста, то есть его семантическое ядро, и статистические показатели. К статистическим показателям относятся: количество символов с пробелами и без, количество слов, в том числе уникальных и значимых, стоп-слов, количество «воды», грамматических ошибок, процент классической и академической тошноты, семантическое ядро. При подсчете учитывается число уникальных слов (без повторений), число значимых слов (существительных), стоп-слов (которые лишены своего смысла).

Главная характеристика, которая интересовала в данной работе - частотность слов. По результатам каждой загрузки корпуса текстов фиксировались данные о частности слов и заносились в таблицу. После чего, на основе полученных данных строились графики частотности упоминания слов в обращениях для выбранных площадок.

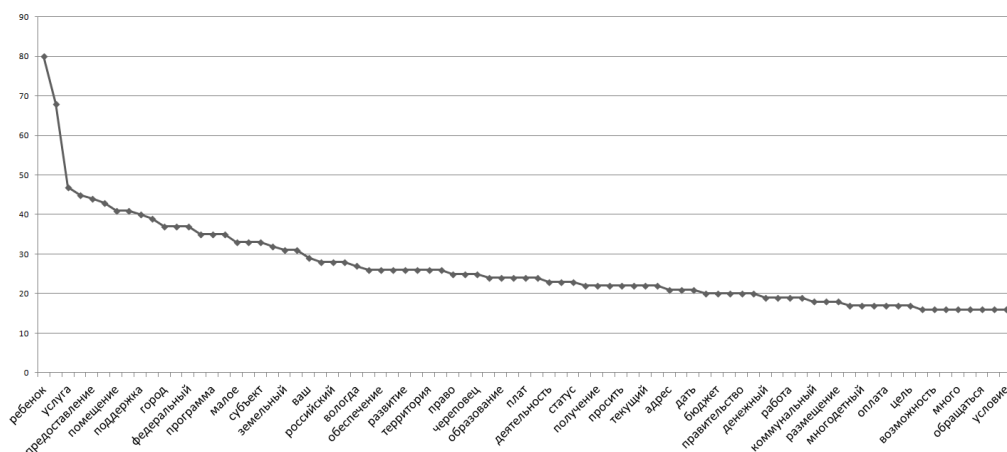


Рис. 1. Результаты семантического анализа портала https://vologda-oblast.ru/onlayn_priyemnaya_gubernatora/

создание кода, поскольку предоставляет набор готовых классов, функций или структур для работы с имеющимися данными.

Однако, порталы интернет приемных органов власти, как правило, не имеют подобных интерфейсов, поэтому для автоматизированного сбора сообщений было решено использовать парсинг данных. Парсинг сайтов является эффективным решением автоматизации сбора и обработки информации. Парсинг – это принятое в информатике определение последовательного синтаксического анализа информации, размещённой на интернет-страницах. Парсер – это программа или скрипт, позволяющая выполнить такой анализ и представить результат в нужном для пользователя виде.

Процесс парсинга html - страницы можно разделить на три основных этапа вне зависимости от языка, на котором он написан. Получение исходного кода html - страницы. Следующий шаг - извлечение из полученного кода нужной информации. Получив исходный код html - страницы, необходимо выполнить над ним обработку, т.е. отделить искомый текст от гипертекстовой разметки, выстроить иерархическое дерево элементов документа (DOM) и извлечь из страницы искомую информацию. Последний шаг – сохранение результата.

При данном подходе парсер нужно настраивать для каждой площадки. В пилотном исследовании парсер был настроен на одну площадку, с которой удалось выгрузить 457 обращений. Время сканирования составило 53 минуты. Для выгрузки были сформированы адреса парсинга путем присвоения каждой ссылке номера страницы, далее парсер осуществил обход по всем страницам, копирую информацию по заданным параметрам. Данный подход позволил сократить время сбора сообщений, при уменьшении трудозатрат.

Для дальнейшей оптимизации планируется использование планировщика заданий, для того чтобы коллекция обращений пополнялась по расписанию, с помощью технологии Cron (планировщик заданий, который работает на Unix/Linux хостинге). Эта технология позволяет автоматически выполнять определенные действия на сервере (запуск программ, скриптов и т.д.), с заданным временем или периодичностью. Задания для Cron можно описать так — это несколько строк (одна строка - одно задание), в которых указывается периодичность запуска и команда (которая означает, какую программу нужно запустить).

4. Заключение

Полученные данные из интернет-приемных показали, что на основе полнотекстовых обращений можно проводить контент анализ. Понимая определенные ограничения представленного исследования, исследовательская группа собирается продолжить работу в данном направлении.

В результате исследования можно сделать заключение о том, что, используя данную методику анализа обращений горожан в органы власти, можно выявлять самые актуальные темы в регионе или муниципалитете. Метод показал себя как эффективный инструмент поиска основных проблем. Однако методика и инструментарий исследования требует дальнейшей доработки. В планах работ автоматизированный сбор обращений со всех площадок, а также обработка информации с помощью более продвинутых систем контент-анализа. Также планируется сопоставление полученных результатов с официальной статистикой интернет-приемных, которую органы власти обязаны публиковать ежегодно.

В последующих исследованиях это позволит оценить, как отсутствие или наличие опубликованных решений может влиять на вовлеченность граждан в электронное участие. Это в свою очередь позволит осуществить анализ социальной эффективности их функционирования. Вероятно, публикация решений может влиять на интерес граждан к использованию данного канала взаимодействия с органами власти.

Работа выполнена при поддержке РФФ, проект №18-18-00360 «Электронное участие как фактор динамики политического процесса и процесса принятия государственных решений».

Литература

- [1] Видясова Л.А., Мисников Ю.Г. Критерии оценки социальной эффективности порталов электронного участия в России // Информационные ресурсы России 2017. № 5 (159). С. 16-19.
- [2] Chugunov A.V., Kabanov Y. Evaluating E-Participation Institutional Design. A Pilot Study of Regional Platforms in Russia // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) LNCS. Springer, 2018. Vol. 11021. P. 13-25. DOI: 10.1007/978-3-319-98578-7_2.
- [3] Литвинцева Е.А. Вовлеченность граждан в процессы государственного управления и контроля // Среднерусский вестник общественных наук, 2017. № 6. С. 94-102.
- [4] Видясова (Бершадская) Л.А., Новиков Д.М. Влияние сообществ в социальных сетях на публикацию электронных петиций и их распространение // Сборник тезисов XX Объединенной научной конференция «Интернет и современное общество» (IMS 2017). С. 66-73.
- [5] Мисников Ю.Г., Филатова О.Г., Чугунов А.В. Электронное взаимодействие власти и общества: направления и методы исследований // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Гуманитарные и общественные науки, 2016. №1. С. 52-60.
- [6] Федеральный закон «О порядке рассмотрения обращений граждан Российской Федерации» от 02.05.2006 N 59-ФЗ (последняя редакция) // URL: http://www.consultant.ru/document/cons_doc_LAW_59999/ (дата обращения: 10.03.2019).
- [7] Типовой общероссийский тематический классификатор обращений граждан, организаций и общественных объединений // URL: <http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=567595#004606215920240908/> (дата обращения: 12.03.2019).

Research of the Russian E-Participation Portals: Case of the Data Extraction from Open Sites

B.A. Nizomutdinov, A.S. Tropnikov

ITMO University

E-participation is becoming an increasingly important factor in the development of successful relations between the state and society. There are different ways to participate electronically. One of them is an appeal to the authorities through the Internet reception. The results of the study show that in Russia there are platforms for submitting citizens' appeals to the authorities, where appeals and answers to them are published in the open access.

The article presents the results of search sites, as well as content analysis of published appeals. The authors have collected a list of Internet reception rooms, where the texts of appeals are publicly available - this is an appeal of citizens with answers. 5 sites from different regions with the maximum number of applications were selected from the general list, on the basis of which the content analysis of published applications was carried out. The proposed findings may be useful for selecting the most appropriate methodology for assessing e-participation in Russia.

Keywords: electronic participation, Internet reception, applications, personal data

Reference for citation: Nizomutdinov B.A., Tropnikov A.S. Research of the Russian E-Participation Portals: Case of the Data Extraction from Open Sites // The State and Citizens in the Electronic Environment. Vol. 3 (Proceedings of the XXII International Joint Scientific Conference

«Internet and Modern Society», IMS-2019, St. Petersburg, June 19-22, 2019). – St. Petersburg: ITMO University, 2019. P. 155 – 163. DOI: 10.17586/2541-979X-2019-3-155-163

Reference

- [1] Vidyasova L.A., Misnikov Yu.G. Kriterii ocenki social'noj effektivnosti portalov elektronnoho uchastiya v Rossii// Informacionnye resursy Rossii 2017. N 5 (159). P. 16-19.
- [2] Andrei V. Chugunov, Yury Kabanov. Evaluating E-Participation Institutional Design. A Pilot Study of Regional Platforms in Russia // International Conference on Electronic Participation ePart 2018: Electronic Participation. P. 13-25.
- [3] Litvinceva E.A. Vovlechnost' grazhdan v processy gosudarstvennogo upravleniya i kontrolya // Srednerusskij vestnik obshchestvennyh nauk, 2017. N 6. P. 94-102.
- [4] Vidyasova (Bershadskaya) L.A., Novikov D.M. Vliyanie soobshchestv v social'nyh setyah na publikaciyu elektronnyh peticij i ih rasprostranenie // Sbornik tezisov XX Ob"edinennoj nauchnoj konferencii «Internet i sovremennoe obshchestvo» (IMS 2017).
- [5] Misnikov Y.G., Filatova O.G, Chugunov A.V. Elektronnoe vzaimodejstvie vlasti i obshchestva: napravleniya i metody issledovanij // Nauchno-tehnicheskie vedomosti Sankt-Peterburgskogo Gosudarstvennogo Politehnicheskogo Universiteta. Gumanitarnye i obshchestvennye nauki, 2016. N 1. P. 52-60.
- [6] Federal'nyj zakon "O poryadke rassmotreniya obrashchenij grazhdan Rossijskoj Federacii" ot 02.05.2006 N 59-FZ (poslednyaya redakciya) // URL: http://www.consultant.ru/document/cons_doc_LAW_59999/.
- [7] Tipovoj obshcherossijskij tematiceskij klassifikator obrashchenij grazhdan, organizacij i obshchestvennyh ob"edinenij // URL: <http://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=EXP&n=567595#004606215920240908/>.