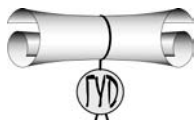


## Трансформация электронных архивов документов и данных на основе краудсорсинга

Коробова Анна Николаевна, начальник Отдела  
службы «одного окна» Префектуры Зеленоградского  
административного округа города Москвы, к.э.н.

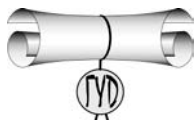
Термин **краудсорсинг** (от англ. crowd - толпа и sourcing - источник, использование ресурсов) означает передачу отдельных производственных функций неопределённому кругу лиц на основании публичной оферты, не подразумевающей заключение трудового договора. При этом используется коллективный интеллект и синергия взаимодействия большого количества людей. Краудсорсинг позволяет агрегировать информацию, опыт, мнения, прогнозы, предпочтения и оценки. Классификация технологий краудсорсинга проводится по функциям и типам решаемых им задач, кругу задействованных участников, типам и способам привлечения людей, способам комбинирования и распределения задач, способам оценки выполнения задач, степени вовлечения компьютера в работу людей, роли человека в процессе краудсорсинга, типам мотивации, способам организации и координации работы, способам агрегации информации и т.д.

Так, например, в зависимости от используемой **технологии сетевых действий** можно выделить **краудголосование**, которое предполагает простое голосование за различные варианты предлагаемых решений, без объяснения своего выбора и предложения других вариантов. **Краудсторминг** предполагает комментирование решений, генерацию идей. Возможно их сочетание, например, когда для краудсторминга используется определенный пул экспертов, а для краудвоутинга – неограниченный круг. **Краудслаппинг** заключается в выплескивании негативных эмоций, возникающих по поводу неправильных действий компании или определенного лица («выпустить пар»), но при этом без предложения чего-либо конкретного. **Краудпроизводство** направлено на создание продукта, имеющего некий сетевой вес, например, производство рекламного видеоролика, программных продуктов, распределенный сбор информации, на основе которых принимаются управленческие решения. В зависимости от привлекаемых участников, возможен **ограниченный краудсорсинг**, который применяется для создания определенного круга людей, имеющих специальные навыки по данному вопросу, т.е. создается некий фильтр, для увеличения КПД предложений людей. **Неограниченный краудсорсинг**, напротив,



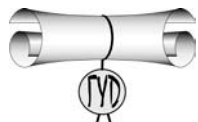
является более незащищенным от непрофессиональных действий некоторых пользователей, и поэтому менее эффективен для оптимизации бизнес-процессов. Однако при этом можно привлечь оригинальные идеи в результате открытого мозгового штурма и нестандартного, междисциплинарного подхода. Четыре года назад автор самого термина краудсорсинг Дж.Хау выделил всего четыре класса краудсорсинга: краудмудрость, краудпроизводство, краудголосование, краудфандинг. С.Карелов, основатель краудсорсинговой компании Витология, отмечает, что в 2012 году даже самые короткие классификации краудсорсинга уже насчитывают десяток классов и подклассов, однако они уже не могут учесть все разнообразие и темпы развития краудсорсинга.

Мотивацией участия в проектах краудсорсинга, краудсторминга и краудпроизводственных инициативах является самореализация, лидерство, признание со стороны коллег, желание улучшить качество товаров, работ, услуг. Материальное вознаграждение при этом или вообще не предполагается или принимает форму участия в доходах, фиксированного приза за лучшие идеи или разовые премии в соответствии со значимостью предложений членов сообщества. Возможна также реальная и/или виртуальная монетизация активности участников на специализированных биржах идей с покупкой/продажей акций идеи, с помощью которой оценивается ее потенциал. Также учитывается мотивационный фактор использования этого механизма для формирования кадрового резерва, социального лифта для талантливых и креативных авторов. Выработанные с помощью краудсорсинга инновационные решения компаний более привлекательны для потребителей, если при их создании были задействованы непосредственно сами потребители, и они могут свободно обмениваться своими идеями о продукте (мотивационный фактор участия в управлении). В результате постепенно стираются грани между устоявшимися векторами информационных потоков в сторону открытых консультаций, разделенного знания, взаимного признания опыта, поиска путей взаимодействия между компанией, сотрудниками, потребителями и поставщиками. Использование краудсорсинга основано на законе Джоя, по которому в любой области деятельности большая часть знания находится за пределами любой действующей в этой области организации. Закон Джоя подтверждается исследованиями Ф.Хайека и Е.Хиппеля, утверждавшими, что знание о продукте распределено среди его пользователей в обществе и принимает форму цены на него в процессе коммуникации продавца и покупателя. Основная задача краудсорсинга, таким образом, состоит в том, чтобы найти доступ к этому знанию.



В России активно развиваются инициативы использующие технологии краудсорсинга, позволяющие привлечь экспертное сообщество для выработки конструктивных предложений модернизации экономики и стимулирования инноваций. Например, в российском проекте «Открытое Правительство» включены механизмы социальных сетей для общественного обсуждения законопроектов, сбора предложений совершенствования системы управления ([большоеправительство.рф](http://большоеправительство.рф), [россиябездураков.рф](http://россиябездураков.рф), <http://zakonoproekt2012.ru/>). Другим примером является недавно запущенный Агентством стратегических инициатив (АСИ) и Witology проект по улучшению инвестиционного климата (сайт проекта <http://asi12.ru>), включивший к настоящему времени коллективную проработку десяти стратегических инициатив от оптимизации условий для экспортной деятельности до реагирования на обращения предпринимателей и налогового администрирования. Лучшие предложения обсуждаются на рабочих группах Национальной предпринимательской инициативы (НПИ), включаются в итоговые документы «дорожных карт», представляемые на рассмотрение Наблюдательному совету АСИ при Правительстве РФ и утверждаемые распоряжениями Правительства РФ. Профессиональная сеть Госбук, Регионалочка, многие аналогичные региональные и муниципальные проекты также основаны на использовании краудсорсинга. Широко распространены такие проекты с использованием геоинформационных систем для обратной связи с правительствами. Например, в Европейских странах люди могут указать местоположение на карте, написать о правонарушении, проблеме и проголосовать за нее. Это пример краудсорсинга для поддержания системной обратной связи и контроля работы подрядных организаций. Подобный проект был запущен в 2011 году Правительством Москвы для оценки состояния и качества уборки дорог, работы и расположения светофоров, дорожных знаков (портал [www.doroga.mos.ru](http://www.doroga.mos.ru)), работ по реализации программ комплексного развития (портал «Наш город» [gorod.mos.ru](http://gorod.mos.ru)). В бизнес-среде использование социальных краудсорсинговых сетей также получило широкое распространение для решения задач по всему маркетинговому миксу 5P во многих отраслях от текстильной промышленности, ИТ, банковского сектора, сферы общепита до автомобильной и даже военной отрасли.

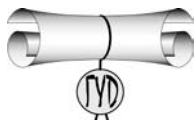
В октябре 2012 года стартовал совместный краудсорсинг-проект префектуры Зеленоградского административного округа и компании Витология по поиску новых путей развития Зеленограда. Специфика этого краудсорсинг-проекта в том, что людьми обсуждается – не как это принято обычно – день сегодняшний, а перспектива, будущее. Участники выявляют проблемы, которые ожидают округ через десятилетие, предлагают



идеи, каким должен быть город в 2022 году, и решения, как этого достичь. Первый месяц работы площадки показал высокий интерес пользователей. Всего зарегистрировалось более 750 участников. Большинство из них зеленоградцы (79%), около 20% – жители России, проживающие как в Москве, так и в других регионах. Есть среди участников и представители зарубежных стран – из Америки, Европы и даже из Сан-Марино. Всем им, как оказалось, интересно поучаствовать в обсуждении будущего интеллектуальной столицы России – Зеленограда. Возрастной диапазон участников колеблется от 14 до 75 лет, хотя опять же преобладает возрастная категория от 20 до 40 лет. На первом этапе участники проекта определяли «тренды», то есть некие тенденции, которые к 2022-му году изменят картину мира. По предположениям краудсорсеров мировыми трендами станут переход к городскому самоуправлению, политика энергосбережения и экологический транспорт. Из российских тенденций были названы отток квалифицированных кадров за границу, приток населения в Москву, увеличение господдержки инновационных предприятий и критическое состояние коммунальных сетей. Зеленоград же на фоне этого грозит увеличение транспортной напряженности, превращение в спальный район, утрата статуса наукограда и отток квалифицированных специалистов в другие округа Москвы и за границу.

В противовес пессимистическим трендам, было следующее задание проекта – его участники должны были ответить на вопрос: "В каком городе вы хотите жить в 2022 году?". Всего было представлено 119 идей будущего Зеленограда от откровенно фантастических до рациональных и вполне воплотимых в жизнь. Из них на следующие этапы генерации условий и мероприятий достижения образов будущего развития Зеленограда, было отобрано четыре идеи: Зеленоград 2022 — один из ведущих российских и известный мировой научно-технологическо-производственный и культурно-образовательный центр (кластер); удобный для жизни город; центр социальных инноваций; центр промышленного туризма.

Кроме того, на площадке развернулись обсуждения по созданию системы велодорожек, превращению зеленоградских библиотек в модное место для коворкинга, высказаны идеи по организации интересного досуга, семейного отдыха в Зеленограде, новых форматах развлечений и новогодних праздников и др. Всего за месяц на площадке сформулировано 455 предложений, оставлено 2828 комментариев, выставлено более 1000 оценок. Проект продлится до конца декабря 2012 года и можно еще успеть изучить на практическом примере как работает эта технология для планирования развития города. Регистрация участников проходит на <http://zelao-prom.ru/future/stat-uchastnikom/> .

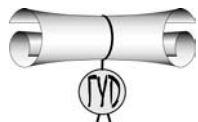


Эти и многие другие примеры показывают, что с помощью сетевого общения можно эффективно решать задачи любой сложности.

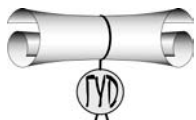
**Как применяется данная технология в работе с архивами?** Здесь выделяются три направления – обсуждение совершенствования работы архивных органов, привлечение «толпы» к оцифровке архивов, «открытые данные».

Во-первых, совершенствование деятельности архивов с помощью краудсорсинга может помочь в дизайне электронных сервисов запроса документов из государственных архивов, обсуждении различного рода документов (концепции, законов, регламентов, стандартов, инструкций и др.). 28 августа 2012 г. на сайте Федерального архивного агентства было опубликовано обращение к общественности с призывом в срок до 15.09.2012г. в целях качественной подготовки Концепции развития архивного дела в РФ, а также учета мнения как можно более широкого круга профессиональной, научной и культурной общественности высказать свои предложения по формату и содержанию указанного документа (<http://archives.ru/press/280812-reference.shtml>). Несомненно, что привлечение экспертного сообщества к обсуждению такого важного документа могла бы иметь при сегодняшнем уровне развития краудсорсинговых технологии более зрелый формат – или уже активно используемый в нашей стране формат обсуждения законопроектов или формат синтеллектуального краудсорсинга или форум с возможностью голосований за поданные предложения и их комментирования. Последний вариант был использован в начале 2012 года Национальным Архивом США, которые открыли специальный форум <http://govrecordmanagement.ideascale.com/> для сбора идей по улучшению управления базой правительственных документов. Всего за период работы проекта на площадку пришло 180 человек, было высказано 30 предложений, к которым поступило 88 комментариев, а также выставлено 273 оценки. Самой комментируемой стала идея про управление информацией в 4-х измерениях (люди, процессы, политики, процедуры). При этом, не столь важно сколько участников пришло и высказало свое мнение на такой площадке, сколько качество высказанных идей.

Вторым направлением краудсорсинга для архивных служб является **привлечение ресурсов «толпы» к переводу архивов в электронный вид**. Здесь используются две различные технологические модели. Во-первых, получила распространение практика привлечения заинтересованных пользователей к проверке, тегированию распознанных текстов, индексированию файлов. Например, в Самаре на основе дистанционной занятости осуществляется перевод в электронный вид документов государственных архивов региона. К оцифровке архива привлечено большое количество людей,



работающих из дома, в том числе лица с ограниченными возможностями, безработные граждане. Аналогичные проекты проводятся или планируются во многих других регионах РФ. Однако в классическом понимании – эти проекты не являются краудсорсинговыми, так как подразумевает заключение контракта и оплату. Существует и другой опыт. Так, **библиотека Конгресса США**, одна из крупнейших библиотек мира, попросила пользователей популярного фотосервиса Flickr идентифицировать людей на фотографиях, которых самостоятельно у сотрудников библиотеки опознать не получилось. Другой пример относится к скандинавским странам. **Национальная библиотека Финляндии** объявила о запуске национальной электронной программы по оцифровке исторического наследия страны. Электронная программа «цифровые волонтеры» (Digitalkoot) использует краудсорсинг, мобилизуя людские ресурсы, чтобы оцифровать миллионы страниц архивных материалов. Автоматизированная платформа разделяет скучные рутинные задачи на серию микрозадач и рассылает их по Интернету. Как только эти микрозадачи выполняются заинтересованными «микроработниками» в Финляндии или за её пределами, провайдер программы Microtask сводит полученные результаты воедино, как кусочки мозаики. Программа в настоящее время состоит из двух онлайн-игр. В первой участникам предлагается два разных слова, и они должны как можно быстрее определить, одинаковые они или нет. Это помогает выявить ошибочные слова в архивных материалах. Во второй игре участники должны правильно написать слова, появляющиеся на экране. На следующем этапе планируется расширение программы и ее нацеливание на более серьезных любителей истории. **Национальная библиотека Австралии** для редактирования и тегирования оцифрованных газетных страниц активно использует краудсорсинг. В этом году волонтеры отредактировали 2 млн 200 тыс. страниц. Таким образом, в этой части работы над оцифровкой возможна экономия с использованием людей, которым интересна история, краеведение или которым просто хочется заниматься общественно полезным делом. Другой современной моделью привлечения людских ресурсов к оцифровке архивов является **использование технологии антиспамовой защиты**. Например, **reCAPTCHA** комбинирует в себе защиту от автоматических регистраций и распознавание текстов. Основным применением технологии стало распознавание текстов отсканированных копий старых книг, газет «The New York Times» и других изданий для сервисов Google Books и Google News. Также недавно корпорация Google начала в тестовом режиме использовать технологию защиты от спама reCAPTCHA для распознавания содержимого табличек на домах, снятых сервисом панорамных снимков Street View. Первоначально технология reCAPTCHA была разработана Университетом



Карнеги-Мемона для защиты сайтов от ботов, а в 2009 году ее приобрела корпорация Google. «reCAPTCHA» используется в продуктах Google, а также в LiveJournal, Twitter и на многих других интернет-площадках. Ежедневно reCAPTCHA демонстрирует пользователям 100 млн. комбинаций.

Третьим направлением совершенствования работы архивов следует выделить публикацию каталогов архивов в формате «открытых данных». Под «открытыми данными» будем понимать информацию, предоставляемую органами власти и публикуемую в специальных машиночитаемых форматах. Кроме того, признаками отнесения опубликованной информации к «открытым данным» является не только техническая открытость, но и правовая, то есть отсутствие ограничений на использование данных по сферам, пользователям и целям, а также на комбинирование с другими данными. Например, недопустимо ограничение «только для некоммерческого использования» или «только в целях образования». Появление «открытых данных» обусловлено, прежде всего, необходимостью создания удобных для пользователя электронных сервисов и приложений на основе принципов частно-государственного партнерства. Связанное различными ограничениями правительство не всегда может реализовать эти возможности из-за дефицита бюджетных средств, необходимости конкурсного размещения заказа, недостатка креатива и др. В архивном деле РФ пока существуют единичные примеры таких массивов. Например, Центр хранения архивного фонда Алтайского края выложил в формате HTML уникальную коллекцию документов 1744-1747 годов - Архив "Комиссии А.В.Безера", ([www.uni-altai.ru/ArchiveBeyer](http://www.uni-altai.ru/ArchiveBeyer)). При активном появлении других массивов «открытых данных» исторических архивов, соответствующей информационной кампании и стимуляции независимых разработчиков в виде, например, конкурсов, как показывает международный опыт, из таких массивов могут появиться новые полезные приложения.

Таким образом, для совершенствования деятельности государственных архивов могут найти широкое применение различные краудсорсинговые технологии, которые будут способствовать экономии бюджетных средств на оцифровку архивов, появлению новых удобных для пользователей электронных услуг и приложений на основе архивной информации.