

Открытые данные и НКО

Использование современных информационных технологий стало привычным. Технологии меняют социум. Выход нового технического устройства на уровень массового потребления, повсеместное использование новой социальной технологии меняет поведение людей, меняет структуру общественных проблем.

Объём ежегодно генерируемой цифровой информации уже перевалил за зеттабайт (примерно 10 в 21 степени байт), т.е. за 1 час создаётся столько же информации, сколько хранится в Библиотеке Конгресса США.

Феномен открытости данных стал возможен благодаря развитию информационных технологий, которые существенно облегчили доступ к Интернет-ресурсам широкого круга пользователей. Концепция открытых данных выходит далеко за рамки прозрачности и отчётности общественных структур. Открытые данные могут превратиться в инструмент социальных изменений. Это открывает новые возможности использования открытых данных в интересах некоммерческих организаций. Однако, как пишет Дэн Сатч (Nominet Trust) в предисловии к обзору под названием «Открытые данные и благотворительные организации»¹, понимание, использование и публикация открытых данных в третьем секторе по-прежнему пребывает в младенчестве.

Ниже представлен обзор, подготовленный на основе международных публикаций и зарубежной практики. В нем даётся представление о том, что такое открытые данные, в чём их ценность и влияние на общество, а также как именно открывать данные.

¹ Dan Sutch, Open data and charities

<https://www.nominettrust.org.uk/sites/default/files/Open%20Data%20and%20Charities.pdf>

Концепция открытых данных

Открытые данные — это данные, которые легко доступны, приспособлены для использования посредством компьютера и которые можно свободно использовать и распространять. Формально, открытыми определены следующие виды данных²:

- 1) доступные онлайн;
- 2) представленные в общепринятом машиночитаемом формате;
- 3) не требуют специального разрешения для использования.

Иногда к этому добавляют ещё и «бесплатность»³.

Несмотря на простоту сформулированных требований, за каждым из них стоят вполне определённые, строго формализованные положения⁴.

«Доступные онлайн» практически означает, что доступ к данным не должен зависеть от разрешения какого-либо субъекта. Как в библиотеке, вы не сами выбираете книгу с полки, а зависите от сотрудника, который эту книгу вам приносит.

«Машиночитаемый формат» означает, что не нужно специально придумывать средства при использовании данных, достаточно воспользоваться некоторой общедоступной компьютерной программой. Тут необходимо дополнить: именно компьютер должен уметь использовать данные. Такие распространённые форматы, как рисунки (графические файлы) или pdf-файлы для этого не подходят. Эти форматы удобны, чтобы показать данные пользователю на экране компьютера, т.е. для визуализации, но для их обработки требуется применение специальных алгоритмов и программ.

Пункт **«Специальное разрешение для использования»** имеет свои нюансы. Полная свобода пользования открытыми данными предоставляется т.н. лицензией Creative Commons. Но она не единственная. Другие варианты налагают на пользователя ограничения, связанные с правом распространять данные или использовать их только для своих нужд или распространять преобразованные данные.

² Hall, W. et al. Open data and charities. Nominet Trust.

www.nominettrust.org.uk/sites/default/files/Open%20Data%20and%20Charities.pdf. Данный доклад часто используется в обзоре уже без специальных отсылок.

³ Open data and the economy – value creation and government’s role and opportunity. Discussion document, November, 2014. McKinsey & Company, 2014. URL: ac.gov.ru/files/attachment/4349.pdf

⁴ Подробнее см., например: www.infoculture.ru/glossary/open-data/; opendefinition.org/licenses/; opendefinition.org/od/1.1/ru/

Что значит – открыть данные?

Открытие данных – не такой простой процесс, как может показаться. Выложить данные в Интернет — нетрудно и доступно почти любой организации, но это не делает данные открытыми.

А если вы всерьёз задумаетесь над тем, чтобы сделать это с максимальным социальным эффектом, то потребуется профессиональная помощь. Как минимум, нужно будет определиться со структурой и форматом, переструктурировать и переформатировать имеющиеся данные.

А чтобы улучшить вторичное использование данных, сделать их автоматически обнаруживаемыми и обрабатываемыми другими приложениями, потребуется привлечь веб-программиста, что представляет трудность для небольших организаций. В этой связи сейчас появляются не только программные продукты, облегчающие этот процесс, но и специализированные НКО и коммерческие организации, предоставляющие такой сервис.

Некоторые рекомендации для НКО⁵:

- Понятие открытости организации не тождественно публикации ею открытых данных. Многие НКО публикуют в Интернете много важной информации, но это не открытые данные. Например, часто публикуются отчёты и доклады, для подготовки которых использовались данные. Иногда в них вставляются и сами данные, но это нужно делать в табличном виде, хорошо и понятно структурированными, что сделает их удобными для машинной обработки. Рекомендуется также всегда публиковать все исходные данные.
- Машиночитаемый формат (наиболее часто употребляемые — CSV, XML и JSON). Любое обновление и дополнение данных не должно приводить к тому, что они становятся несовместимыми с предыдущей версией и пользователю придётся приспосабливаться к новой организации, новой структуре данных.
- Используемый формат должен быть понятен и человеку, чтобы при вторичном использовании данных не нужно было ломать голову и выяснять, какой формат использован. С этой целью предлагается применять «открытый формат данных», который стандартизован и при необходимости можно без труда найти описание.
- Если открытые данные публикуются не на специально отведённом сайте, то их лучше все сконцентрировать в папке, путь к которой начинается с имени /open (косая черта Open).

⁵ Andy Isaacson A. Open data, done right: five guidelines. Markets For Good. URL: <http://www.marketsforgood.org/beyond-alphabet-soup-5-guidelines-for-data-sharing/>

- Самый продвинутый подход подразумевает использование программного кода API (application program interface⁶), который обеспечивает взаимодействие двух компьютерных программ между собой.

В сети можно найти много руководств, примеров и рекомендаций. Например, Фонд «Открытые знания»⁷ (Open Knowledge Foundation) системно занимается продвижением тематики открытых данных (в широком смысле), на его сайте⁸ можно найти много полезных ресурсов.

Собственно, планируя сбор данных, уже нужно знать – собираетесь ли вы делать их открытыми? И если да, то предусматривать это в последующих процедурах.

Организации, открывающие данные, сталкиваются с рядом **технологических** проблем:

1. Использование подходящих инструментов хранения, доступа и обработки данных. Большие массивы данных или сложно структурированные наборы неудобно обрабатывать такими популярными офисными программами, как Word и Excel.
2. Качество описания данных и проверка источников требует специальных усилий и опыта. В противном случае конечный пользователь будет получать не те данные, которые ожидал (иногда даже не заметив этого).
3. Качество данных и управление данными – любая публикация данных подразумевает ответственность.

Открытие данных предъявляет повышенные требования к организации, публикующей открытые данные, а соответственно, сопряжено с рядом **организационных** сложностей:

1. Культура данных: специфическое отношение к данным как к важному и реально используемому ресурсу, а не побочному или вспомогательному эффекту деятельности.
2. Организационное развитие и грамотность: уровень развития организации можно измерять по её работе с данными. Не нужно думать, что все сотрудники готовы качественно работать с данными.
3. Стратегия, руководство и менеджмент: открытие данных — это не просто дополнительная услуга обществу, а отдельное направление деятельности, согласующееся с миссией организации.
4. Управление процессом открытия данных должно быть продумано, спланировано, обеспечено всеми необходимыми ресурсами.

⁶ <http://searchexchange.techtarget.com/definition/application-program-interface>

⁷ Руководство по открытым данным. Open Knowledge. URL: <http://opendatahandbook.org/guide/ru/>

⁸ <https://okfn.org/>

5. Лицензирование: чёткое понимание того, насколько правомерно использование чужих данных и обозначение прав на использование публикуемых вами данных, особенно, если имеются ограничения.
6. Использование и предоставление данных возлагает дополнительную ответственность. Открывая данные, вы становитесь субъектом дополнительных правовых отношений.

Отдельные темы — обеспечение юридической защиты данных, защиты персональных данных и техническая защита данных – требуют технической и организационной готовности и зрелости НКО.

Эти сложности могут смутить тех, кто не привык работать с данными. Тем не менее следование несложным рекомендациям необходимо, чтобы Интернет не превратился из полезного хранилища данных в их «помойку».

Тим Бернерс-Ли предложил шкалу оценки открытости данных⁹. Выложенные в интернете массивы оцениваются по 5-балльной шкале — от просто, кое-как выложенных, но вполне пригодных для свободного использования данных до систем, которые позволяют другим пользователям подключать свои данные так, что они становятся частью общего массива с удобным интерфейсом для пользователя. Характерно, что предложенные качественные характеристики совсем не подразумевают, что надо стремиться к наивысшей (пятой) степени открытости. У всех степеней есть свои плюсы и минусы. Создаваемая система должна быть такой, чтобы эффективно решать те задачи, которые ей адресованы. Т.е. нужно выбирать не самый развитый, а самый подходящий Вам инструмент.

Насколько открыты данные и насколько значима ценность открытости?

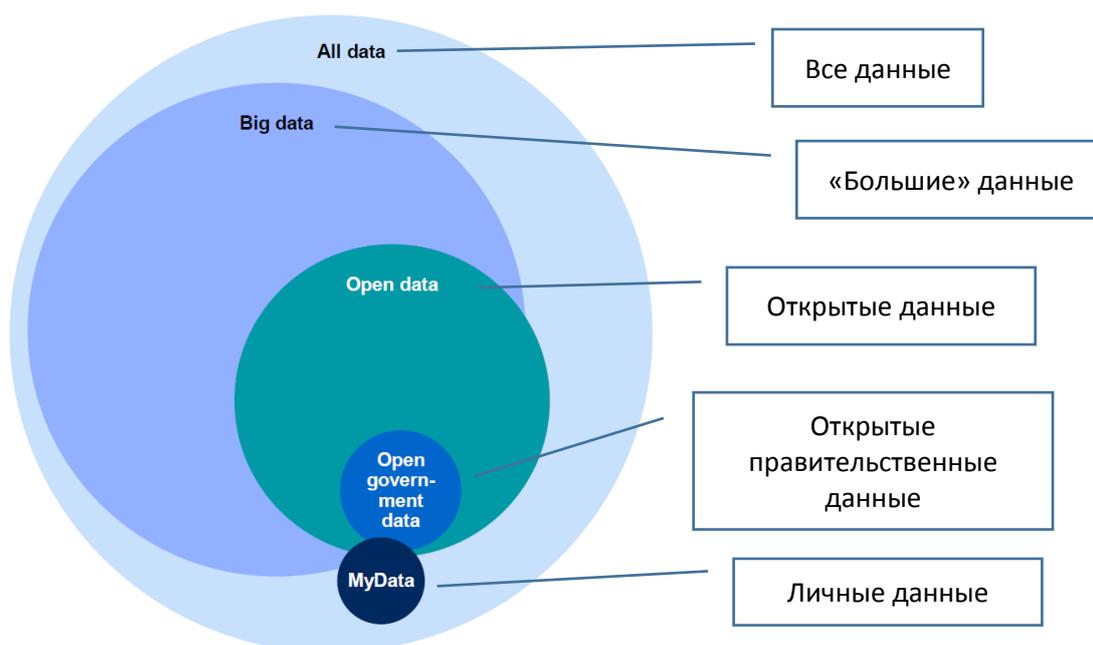
Нельзя сказать, чтобы данные мало открывались или организации не пользовались открытыми данными. Примеров использования открытых данных множество. Но большинство авторов нет-нет, да и упоминает, что мы далеко не исчерпываем имеющийся потенциал. Не будет преувеличением сказать, что мир был бы другим, если бы последние десять лет все, кто контролирует хоть какие-то данные, сделали бы их открытыми, легко доступными для других.

В некоммерческом секторе картина гораздо беднее, чем в других, что отмечается многими наблюдателями и исследователями¹⁰.

⁹ <http://5stardata.info/en/>. Сайт будет полезен тем, кто серьёзно занимается темой «открытые данные».

¹⁰ <http://www.idealware.org/>. Idealware проводил исследования, как НКО используют данные. Согласно полученным результатам, анализ данных ещё не стал общепринятым инструментом повышения эффективности деятельности НКО, несмотря на компьютерную оснащённость третьего сектора и тот факт, что НКО реально собирают очень много данных. Пока НКО не научились делать так, чтобы данные «работали» на результат.

На рисунке ниже показано соотношение всех доступных данных¹¹, согласно исследованию МакКинзи.



При этом правительства более, чем 40 стран мира создали специализированные порталы открытых данных. Только в США насчитывается 90 тысяч массивов открытых данных, опубликованных органами власти, а во всех странах их насчитывается более миллиона.

Как бы ни были велики эти показатели, они показывают, что «открытые данные» не так уж и широко используются. Например, открытые данные публикуются не более, чем пятой частью национальных правительств, а посещаемость таких сайтов в разы уступает количеству просмотров выложенных в Интернете видео.

Данные по некоммерческому сектору отсутствуют, но они составляют мизер от того, что делают правительства.

Оценки экономического эффекта открытия данных¹², видимо, не сильно вдохновляют держателей данных. Стимулы открытия данных проявляются слабо в обществах, где субъекты не испытывают достаточного внешнего «давления», заставляющего их действовать более эффективно. Это относится и к структурам власти (в развитых демократиях они испытывают постоянный общественный контроль и сильное политическое давление), и к некоммерческим организациям (контроль доноров, включая органы власти, независимых общественных групп). И это относится не только к

¹¹ Manyika, J. et al. Open data: Unlocking innovation and performance with liquid information. октябрь, 2013 McKinsey & Company, 2014. http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information

¹² Не очень строгий подсчёт показывает, что ежегодная экономия по всему миру составляет 3-5 триллионов долларов США.

открытости данных, но и к любым новым технологиям, инновациям. «Открытые данные» — это один из приёмов повышения эффективности деятельности за счёт применения современных информационных технологий.

Основные преимущества, которые создаются благодаря использованию открытых данных¹³:

1. Обеспечение большей информированности и прозрачности, лучшее понимание возможностей выбора для индивидуальных потребителей и компаний, что затрудняет мошенничество и снижает риски.
2. Обеспечение сведений для сравнения при тестировании и апробации новых возможностей и формирования стандартов.
3. Адресное обеспечение потребителей за счёт более точного понимания их предпочтений и сегментирования.
4. Повышение оперативности и качества, расширение выбора и углубление понимания в процессах принятия решений.
5. Содействие инновациям и оптимизации процессов за счёт улучшения оценки вариантов.

Для некоммерческих организаций вполне подойдут следующие аргументы в пользу открытых данных¹⁴:

1. Лучшее понимание, что происходит, каковы потребности сообщества.
2. Повышение эффективности деятельности.
3. Улучшение понимания социальных результатов и удалённых эффектов.

К этому можно добавить ещё некоторые соображения относительно важности открытия данных в современном мире.

Первое связано с такими понятиями, как «экономика знаний» и «креативная экономика». В наши дни знания и идеи стали предметом рыночных отношений. Ещё недавно заинтересованные лица гонялись за информацией — чем больше знаешь, тем имеешь более выигрышное положение в любом деле. В значительной степени это справедливо и сейчас. И у людей, и у корпораций есть немало оснований скрывать данные, не делиться ими. Но уже сейчас можно видеть смещение: всё большую ценность имеет не информированность, а идеи, понимание того, как решать ту или иную проблему. Т.е. монетизируются не столько сами данные, сколько инновации, синергия и сотрудничество, основанные на доступе к данным. Не договариваясь об открытии данных с другими, проигрывает всё общество и каждый в отдельности.

¹³ Open data and the economy – value creation and government’s role and opportunity. Discussion document, November, 2014. McKinsey & Company, 2014. URL: ac.gov.ru/files/attachment/4349.pdf

¹⁴ Lucy de Las Casas, Tracey Gyateng and David Pritchard. The Power of Data: Is The Charity Sector Ready to Plug In. NPC. <http://www.thinknpc.org/publications/the-power-of-data/>

Выгоднее открывать данные в свободный доступ. Это работает ещё и потому, что существует много источников данных, и они в любом случае станут общим достоянием, тогда как сокрытие данных не даёт преимуществ. Гораздо лучше открывать данные осмысленно, чтобы иметь возможность извлекать преимущества из этого, а не из того, что вы знаете больше, чем другие.

Расширение общественного движения за открытость данных представляет собой коалицию не только гражданских и некоммерческих организаций, оно получает поддержку благодаря тому, что в этом всё больше заинтересованы рыночные структуры, которые «делают деньги» на открытых данных.

Открытые данные и сообщество

Развитие информационного пространства связано не только с насыщением его данными и массовостью доступа. Наиболее важным фактором развития становится то, что каждый участник становится источником данных. «Сеть пользователей» уже давно превратилась в «сеть данных». Именно в этом суть того, что именуется Web 2.0. Получается, что чем больше участников, тем больше данных, в свою очередь это привлекает ещё больше участников. По известному закону Мэткалфа^{15,16}, полезность сети пропорциональна квадрату количества участников этой сети. Сейчас отмечается новый этап развития: создаваемое для Интернета программное обеспечение автоматически организует любую информацию, поступающую в сеть, в структурированные данные. Имеющиеся собрания текстов — это уже не просто «куча контента». Любое сообщение (например, в Википедии) включается в базу данных и его можно легко подключить к другим ресурсам. Т.е. входящий поток текстов превращается в открытые данные независимо от участников.

Объединение индивидуальных массивов данных в общий массив при правильном подходе даёт синергетический эффект — каждый участник получает от объединённого массива намного больше того, что в него вкладывает. Тем не менее, само собой это не получается. Создание объединённого массива нужно планировать, о нём нужно договариваться, его нужно организовывать и нужно достичь «критической массы», после чего наращивание полезности происходит автоматически и с лихвой оправдывает начальные вложения.

Создание объединённых систем открытых данных — особая задача, которая комбинирует решение технических вопросов и решение задач консолидации сообщества. И эти задачи неразрывны. Если есть сообщество, то можно ожидать успеха

¹⁵ Briscoe, B., Odlyzko, A., and Tilly, B. Metcalfe's Law is Wrong. IEEE Spectrum, 2006. URL: spectrum.ieee.org/computing/networks/metcalfe-law-is-wrong

¹⁶ Weinman, J. What If Metcalfe's Law Is Wrong? Gigaom, 2010. URL: gigaom.com/2010/03/12/what-if-metcalfe%E2%80%99s-law-is-wrong/

(в том смысле, что созданная система будет генерировать высокую общественную полезность данных) в построении общей системы данных.

С другой стороны, наличие хорошего открытого массива данных в значительной мере способствует формированию сообщества, т.е. общественная значимость информационного ресурса привлекает больше участников. Чтобы правильно спланировать оба процесса, и чтобы они дополняли друг друга, нужно специально анализировать ситуацию.

Систему открытых данных сравнивают с экосистемой. Здесь тоже много разнородных игроков со своими запросами, интересами, возможностями и связями между ними. Задача открытия данных не сводится к выкладыванию данных на некоторый сайт в Интернете. Задача – смоделировать сначала саму экосистему (в частности, «просчитать» сетевой эффект) и на основе этой модели решать все технические вопросы. Только в этом случае данные действительно работают на сообщество, генерируют добавленную социальную ценность.

Особо отмечается роль стандартов. Они важны при выборе форматов данных и их описаний. Следование стандартам сокращает путь пользователя к опубликованным данным, не требует много времени, чтобы разобраться, как устроены данные и как их использовать имеющимися в его распоряжении средствами.

При этом по многим вопросам стандартов ещё не выработано, что осложняет задачу.

Как и всякое другое общественное явление, открытые данные несут в себе некоторые угрозы и риски для сообщества. Это особенно проявляется сейчас, когда ещё не накоплено достаточного опыта для регулирования и стандартизации. Безусловно, самым критичным и постоянно будоражающим общественное мнение вопросом представляется баланс использования доступных данных в интересах обеспечения общественной безопасности и одновременное ограничение в интересах защиты личности. Пока продвижение идет от скандала к скандалу в поисках «золотой середины». Скорее всего, спецслужбы будут постепенно расширять свои права использования данных, а для граждан будут вводиться всё новые ограничения.

Основные риски открытых данных для людей заключаются в следующем:

- нарушение права на частную жизнь при открытии персональной информации;
- личная безопасность, связанная с недостаточной защитой данных;
- безопасность в общественных местах при том, что открытые данные могут быть использованы и в антиобщественных целях.

Для организаций выделяют тоже 3 типа рисков:

- нарушение конфиденциальности при действиях в условиях большей прозрачности;

- повышение ответственности и наказания в случае утечек и других нарушений из-за неэффективной защиты данных;
- коммерческие потери при нарушении прав на интеллектуальную собственность из-за отсутствия адекватного регулирования и системы предотвращения незаконного использования данных.

Проблемы здесь только нарастают, заставляя и частных субъектов, и правительства увеличивать расходы на безопасность в виртуальном пространстве.

Открытие правительственных данных для НКО

Данные исследований и данные правительства традиционно рассматриваются как первые кандидаты на открытие. Прежде всего, это связано с тем, что эти данные формируются на средства бюджета, поэтому «принадлежат всем». Сейчас этот постулат подкрепляется экономической выгодой, полезностью для общественного прогресса, потому что в последние годы существенно расширились и удешевились возможности обеспечения доступа к данным, их стало много и проще извлекать из них общественную пользу.

Давление на правительства определяется ещё и требованием прозрачности. Открытие данных — это важный элемент обеспечения прозрачности власти, гарантии обоснованности и эффективности общественной политики. В этом смысле наибольший интерес представляют именно те данные, которые не просто информируют о ситуации, а которые позволяют оценивать деятельность властей. Т.е. выделяются данные «открытого правительства» и «данные, открытые правительством». Первые важны для оценки самого правительства. Вторые — для экономики знаний, как общественный ресурс.

Специальные усилия по открытию данных правительствами — относительно недавняя история. Например, в США и Великобритании специальные порталы появились только в 2009 и 2010 годах.

Например, в Израиле процесс открытости бюджета¹⁷ начался с 2010 года, после перепалки между правительством и пожарной службой. После очередного крупного лесного пожара пожарники объяснили свою беспомощность хроническим недофинансированием. Правительство с этим не согласилось. Общественность не могла разобраться в этом вопросе из-за недоступности чёткой бюджетной информации. Приведение в порядок бюджетной информации только на текущий финансовый год потребовало немало усилий. Но эта открытость тут же выявила, что исходный бюджет вследствие многочисленных поправок за год изменился на 13%. Стало ясно, что с управлением бюджетными ресурсами далеко не всё в порядке. Тогда правительство запросило бюджетную информацию за несколько предыдущих лет. Но тут

¹⁷ <http://opendatahandbook.org/value-stories/en/israel-budget/>

выяснилось, что она доступна только в совершенно неподходящем для анализа формате. Тогда была создана специальная Интернет-платформа, через которую правительство обратилось к гражданам помочь оцифровать бюджетные данные. Всё это привело к появлению в 2014 году хорошего сайта, где визуализированы бюджетные данные.

Нельзя сказать, что бюджетная информация была закрыта для израильтян до этого. Но она была доступна в «неудобном для понимания» виде, поэтому и не обсуждалась. Открытие данных вовлекает граждан в обсуждение и принятие решений на совершенно другом уровне.

История открытия бюджетных данных в Израиле интересна ещё и выводом, который сформулировал главный специалист этого процесса Машон Зер-Авив (Mushon Zer-Aviv): «Чем больше ты расширяешь доступ, тем глубже становится кроличья нора»¹⁸. Открытие данных повышает спрос на новые данные. Открытие данных отвечает на многие вопросы, но ставит ещё больше новых. Планируя открытие данных, нужно отдавать себе отчёт в том, что оно не ограничится одним действием.

Публикация данных в удобном формате позволяет также правительству исправлять ошибки в своих данных. Они обнаруживаются практически всегда, когда открываются какие-либо данные.

4 основные функции власти в содействии открытости данных:

1. *Как субъект политики и законодатель*, власть регулирует использование данных, устанавливает стандарты качества и форматов данных, обеспечивает права собственности и защиту персональных сведений.
2. *Как поставщик данных*, власти собирают информацию, публикуют данные в электронном виде (т.е. делают их открытыми), способствуют повышению качества данных.
3. *Как пользователь*, власть использует открытые данные в процессах принятия решений и отчётности, управлении общественными ресурсами.
4. *Как влиятельный субъект*, власть способствует формированию культуры данных, выступает примером и стимулирует других общественных субъектов.

Нет сомнения в том, что власть — самый крупный потребитель и поставщик общественных данных. Власти постоянно находятся под сильным политическим давлением и открытие данных — очень удобный инструмент обоснования своей политики.

Главный мотив открытия данных органами власти — сокращение бюджетных расходов. Если невозможно помочь обществу путём бюджетного финансирования, то можно предоставлять данные. Это дешевле, но даёт не меньший эффект, если его выразить в деньгах. Бюджетная экономия за счёт открытых данных возможна также потому, что в странах с относительно небольшой долей государства в социальной сфере органы власти выступают не поставщиками услуг населению, а заказчиками. Конечными

¹⁸ <http://opendatahandbook.org/value-stories/en/israel-budget/>

исполнителями при этом по большей части выступают некоммерческие организации. Органы власти заинтересованы снижать цену заказа без снижения качества и объёмов услуг, поэтому они заинтересованы повышать свою информированность (как потребители). В свою очередь, НКО заинтересованы иметь данные для того, чтобы получать государственные заказы. Обе стороны заинтересованы в том, чтобы доказывать свою эффективность.

Вот пример из британской жизни – как данные Министерства юстиции используются некоммерческим сектором. Это тот случай, когда заинтересованы обе стороны¹⁹.

НКО работает с правонарушителями, стремясь снизить уровень рецидивов. НКО предоставляет министерству списки подопечных своих программ, а министерство по своей закрытой базе данных вычисляет некоторые статистические характеристики по выборке, близкой к той, которая указана НКО. Тем самым НКО может сравнить свои результаты со средними показателями по стране.

Открыть сами данные министерство не может, но эти данные важны для вычисления эффективности услуг НКО. Большинство НКО финансируются из бюджета, поэтому заинтересованы обе стороны. Чтобы не предоставлять конфиденциальную информацию, министерство создало отдел — Justice Data Lab — который осуществляет по запросам НКО необходимые вычисления. НКО получают не данные, а результаты расчётов. Сервис бесплатный.

Этот пример можно лишь условно отнести к «открытым данным», скорее – механизм открытых данных заменяется другой услугой, но выполняет почти ту же функцию. Обычно потребителю данных нужны не сами данные, а результаты их некоторого анализа. Пример показывает, что таким образом можно преодолеть ситуацию «мы не можем открыть данные, поскольку в них конфиденциальная персональная информация».

Интересно также, что при коммуникациях с органами власти НКО должны пользоваться специальным адресом электронной почты, который наделён повышенным уровнем безопасности.

При обработке данных главное — статистическая значимость предоставляемой министерством информации, поэтому можно предположить, что в соответствующем отделе министерства работают специалисты по математической статистике.

Одного лишь факта, что открытые данные — это полезно для общества, недостаточно, чтобы данные активно публиковались в удобной форме. Требуется прямая заинтересованность (если не выгода) всех сторон, от которых зависит, что открывать.

Второй аргумент в пользу открытия данных властью сводится к тому, что в демократических обществах уже давно развиваются механизмы гражданского участия в принятии решений. Это стимулируется необходимостью обоснования политики, а также пониманием того, что это действительно улучшает решения, повышает эффективность на этапе внедрения. В этой связи предпринимаются усилия и для того,

¹⁹ <https://www.gov.uk/government/publications/justice-data-lab>

чтобы активная и заинтересованная часть общественности была проинформирована не хуже самого правительства.

Третий мотив — понимание того, что свобода информации — это важный фактор общественного прогресса. Это один из тех случаев, когда права человека влияют на благосостояние. В либеральных общественно-политических системах основная роль власти — создавать условия для того, чтобы общество эффективно справлялось с общественными проблемами без вмешательства правительства. Для этого необходимо повышать информированность, максимально снимать любые ограничения. Ещё недавно распространение общественно важной информации было уделом только СМИ. С развитием Интернета и средств коммуникации инструментарий донесения информации расширился.

Сотрудничество власти с некоммерческим сектором по проблемам открытых данных связано прежде всего с выработкой стандартов качества, доступности, безопасного использования, а также с продвижением и финансированием повышения квалификации различных субъектов, работающих с данными.

Хотя органы власти во всех странах располагают наибольшими объёмами потенциально общественно полезной информации, необходимо задуматься относительно политики (стратегии) открытия этих данных, чего нельзя сделать одномоментно и на что тоже требуются ресурсы. Самое главное, что никто не может дать гарантий, что накопленные в предыдущих политических поколениях данные, достаточно точны. Всегда есть опасение, что публикация этих «выдуманных» данных нанесёт вред обществу²⁰. Ведь, официально опубликованные данные пользуются наибольшим доверием.

Активные и грамотные действия власти по раскрытию данных базируются на комбинации экономической полезности открываемых данных, социальном эффекте и политической мотивации. Это даёт основания для оценки как конкретного массива открытых данных, так и для формирования политики открытия.

В странах с развитой демократией политика и действия власти в любом случае определяются условиями высокой зависимости от групп общественных интересов. Это распространяется и на развитие систем открытых данных. В странах, где механизмы гражданского участия развиты слабо, усилия правительства по открытию данных будут смещены в пользу решения политических задач. В этом случае правительство будет склонно отчитываться «валовыми» показателями — сколько массивов данных и в каких объёмах открыто. Там, где политическая участь сильно зависит от благосостояния граждан, открытие данных строится, исходя из решения общественных проблем, а

²⁰ Достаточно напомнить о фальсификации данных правительством Греции ради получения кредитов. Или многочисленные факты искажения данных голосования в странах со слабыми демократическими институтами. Обнаружение этих нарушений может повлечь за собой целую цепочку неприятностей, выходящих за рамки одной страны.

оценка и планирование открытия данных публичными структурами определяется общественным благом.

В развитых странах приоритет отдаётся тем данным, открытие которых имеет наибольший социально-экономический потенциал. Простые алгоритмы оценки для этого не так сложно разработать. В менее развитых странах правительства публикуют то, что наиболее готово (потому что не сильно затратно) и что не угрожает власти ни политически, ни экономически. В этой связи было бы правильно и честно, на правительственных порталах открытых данных в каталогах указывать планируемые экономические и социальные эффекты от публикации, а затем оценивать реальный эффект.

Но важнейшим источником для формирования таких приоритетов должно быть само общество. Гражданские общественные организации и ассоциации, представляющие интересы общественных групп, могли бы формировать чёткий и обоснованный запрос на то, какие данные следует открывать в первую очередь, которые больше всего помогли бы социальной эффективности соответствующих общественных институтов.

В России правительства и органы власти публикуют много полезной информации, но, как указывают наблюдатели²¹, — это не открытые данные. Компьютерный доступ к ним осложнён.

Официальная политика заслуживает особого рассмотрения и обсуждения. В официальных бумагах и заявлениях можно увидеть, что и как открывается, но с ходу невозможно понять, почему именно это и как принимаются решения по этому вопросу. План и мероприятия понятны, а политика в этой сфере — нет.

Российский эксперт Иван Бегтин²² разъясняет разницу в подходах к построению этой политики. Эффективный подход, когда правительство делает данные открытыми (что не очень дорого), а интерфейс изготавливается частными компаниями. Они делают приложения, благодаря которым открытые данные становятся массово используемыми данными, а потому и их открытость приобретает социальное измерение, полезность. Данные начинают «работать», а не просто «быть открытыми».

Это, кстати, полезное основание для оценки — стремиться надо не к открытости данных, а к тому, чтобы они работали на общество. И оценивать не по количеству и объёмам открытых массивов, а потому, какой общественный эффект, добавленную стоимость они производят.

Другой подход, который характерен для России, - тратить деньги (госзаказ) на производство приложений. Это не даёт возможности развиваться рынку, а бесплатные приложения не используются.

²¹ <http://habrahabr.ru/company/infoculture/blog/198198/>

²² Бегтин И. Приоткрытые данные. Заметки по следам совета по открытым данным 15 июня URL: <http://habrahabr.ru/company/infoculture/blog/260389/>

В целом, можно заключить, что официальные структуры пока специально не озабочены «обслуживанием» некоммерческого сектора. Они открывают данные либо механически, выполняя некоторый план, установленный для ведомства. Либо они стремятся обеспечить как можно больше данных для решения некоторой общественной проблемы, тогда публикуемые данные адресованы всем, кто ею занимается.

Замечено также, что главными пользователями правительственных открытых данных являются сами сотрудники правительственных структур.

Открытый обмен данными между НКО

Для НКО открытые данные — это часть информационного пространства, в котором осуществляется их деятельность. Здесь они выступают и как потребители открытых информационных ресурсов, и как потенциальные поставщики данных в открытый доступ. Независимо от желания организации, они оказываются в новой реальности, избежать которую не удастся. Наиболее продвинутые организации будут активно использовать эти новые возможности и, тем самым, усилят свои лидерские позиции. Отказ от использования открытых данных пока не сильно сказывается на потенциале организации, но очень скоро это будет заметным, и они сами будут ощущать отставание по многим важным параметрам качества услуг и ресурсного обеспечения.

Основной интерес НКО к открытым данным определяется тем, что открытые данные предоставляют дополнительные возможности генерировать социальную ценность (social value).

Другой важный фактор вытекает из характерных особенностей некоммерческого сектора, само существование которого объясняется информационной асимметрией, невозможностью в ряде обстоятельств установить справедливую рыночную цену и ограниченными возможностями вовлечения в принятие решений именно представителей тех социальных групп, ради которых и осуществляются социальные проекты (например, разного рода маргинальные группы). Наполненное информационное пространство позволяет решать именно эти проблемы, открытые данные представляют собой эффективный инструмент, оправдывающий деятельность некоммерческого сектора и обеспечивающий эффективную мобилизацию общественных ресурсов на решение общих проблем.

Отмечается также, что в открытых данных в первую очередь заинтересованы именно небольшие НКО. Возможно, третий сектор более активно продвигал эту тему, если бы все организации были крупными. Крупные организации имеют ресурсы, чтобы самостоятельно обеспечивать себя необходимыми данными и обосновывать свою деятельность перед донорами и обществом. Для небольших организаций, которых большинство даже в тех странах, где некоммерческий сектор считается хорошо развитым, качественная информационная поддержка и обоснование требуют непосильных расходов. Открытые данные компенсируют нехватку ресурсов.

В докладе Nominet Trust²³ приведены 9 аргументов – как НКО может извлечь пользу, воспользовавшись открытыми данными:

1. Формирование стратегии и планирование.
2. Координация деятельности разных организаций.
3. Обеспечение интегрированных сервисов, смежных услуг или пакетов (например, по принципу «два в одном»).
4. Поддержка общественных кампаний, защиты прав, представительства интересов (advocacy).
5. Улучшение оценки удалённых социальных эффектов.
6. Улучшение условий для фандрайзинга.
7. Обеспечение коммуникаций и согласование местных, региональных и общенациональных интересов.
8. Повышение доступности и эффективности управления общественными ресурсами (государственные заказы, субсидии и гранты).
9. Мобилизация сообщества — добровольцев и коммерческих структур.

Нетрудно заметить, что этот список в самом общем виде охватывает все аспекты функционирования и успеха НКО. Можно даже сказать, что открытые данные будут полезны всегда, когда требуется информация — при планировании, внедрении и оценке деятельности НКО, если и то, и другое, и третье выполняется разумно и творчески.

Наибольшая полезность открытых данных связана с **краудсорсингом** — массовым вовлечением граждан в сбор данных. Это особенно характерно для проектов, где результат во многом зависит от оперативности получения данных или потребности с большим объеме данных. Например, при поиске пропавшего ребёнка открытые данные позволяют быстро мобилизовать широкий круг и координировать усилия в реальном времени. В данном случае это является критичным для успеха, используются ранее недоступные преимущества информационных технологий.

Ещё один тип проекта, широко использующий технологию открытых данных, связан с **защитой прав потребителей**. Тут НКО напрямую делают данные открытыми, чтобы выполнять свою миссию. Например, публикация данных о ценах в магазинах микрорайона не только влияет на информированность жителей, но и меняет поведение торговых точек.

Важно, что именно НКО становятся основными потребителями онлайн-платформ, обеспечивающих открытие данных. Как правило, у НКО нет ресурсов, чтобы создавать собственные инструменты, как это делает правительство или бизнес.

²³ Hall, W. et al. Open data and charities. Nominet Trust.
www.nominettrust.org.uk/sites/default/files/Open%20Data%20and%20Charities.pdf

Большое распространение получают сетевые ресурсы, которые объединяют данные из разных источников. Тем самым формируются уже готовые подборки по темам, которые реструктурируются, а для пользователя выглядят как единый набор данных.

Самый простой вариант такого сайта-интегратора есть и в России – «Прозрачность правозащитных организаций»²⁴ (НП «Информационная культура»), содержащий 14 массивов данных об НКО, взаимодействующих с государством. Ресурс представляет собой каталог открытых данных, но не связанных между собой. То, что федеральные структуры расширяют публикацию открытых данных – безусловно, хорошо. Однако по данным не прослеживается ни политики, ни идеи открытости (зачем открывать данные?). Более того, данная инициатива — публикация в одном месте каталога доступных данных — пока не поддержана другими организациями, а сами массивы полезны лишь ограниченному кругу лиц.

Примеры открытых данных, созданных в интересах НКО:

<http://data.foundationcenter.org/> — Центр фондов (США)

<http://www.datakind.org/projects/> — международная команда DataKind (центральный офис - Нью-Йорк) объединяет волонтеров в сфере ИТ, готовых прийти на помощь в конкретных проектах или просто подучить конкретную НКО.

<http://initiatives.marketsforgood.org/> — сайт Markets For Good (США), представлены результаты исследования, проведенного организацией Idealware, специализирующейся на содействии автоматизации и применении информационных технологий среди НКО. Исследование (2012) выявило 309 инициатив создания группами НКО объединённых онлайн-ресурсов по обмену результатами своих проектов. Это базы знаний, реализованных в виде открытых данных.

http://data.gov.uk/sib_knowledge_box/toolkit — пример открытия информации государственным ведомством для государственных служащих. Содержит базу данных типичных цен (вариант прайс-листа), на который могут ориентироваться специалисты, заключающие государственные контракты на выполнение социальных проектов. Такая база данных упрощает процесс анализа бюджета заявки на финансирование.

<http://opende toolkit.net/> — портал открытых данных по международной помощи. Цель – снижение дублирования и повышение кооперации в проектах, направленных на оказание помощи конкретной стране. Любой заинтересованный человек может узнать – кто и чем помогает его стране, кому помогает его страна. По этому portalу видно, как трудно добиться, чтобы все, кто осуществляет программы международной помощи, поставляли свои данные в единую базу — сообщество ещё только наращивается.

<http://ckan.org/> — благотворительный сервис по созданию «открытых данных» CKAN — один из ресурсов Open Knowledge Foundation,

²⁴ <http://data.prongo.ru/>

предоставляющий программное обеспечение, которое можно бесплатно использовать для создания собственного портала.

<http://datahub.io/> — бесплатная онлайн-платформа, помогающая создавать собственный портал, использует программное обеспечение SKAN.

<http://www.socrata.com/> и <http://www.tableau.com/> - коммерческие сервисы.

Развитие культуры данных, открытие данных и их использование — это инвестиции в способности организации выполнять свою миссию, фактор конкурентоспособности в пространстве благотворительности и распределения грантов и заказов, в поддержании высокого уровня методологической обеспеченности и доказательности своей деятельности в общественных интересах.

Развивать культуру данных в организациях помогает профессиональный – то есть предоставляемый специальными внешними организациями – сервис.

В то же время открытие данные несёт некоторые риски, которые нужно принять во внимание. С одной стороны, открытость — это средство установления доверия, с другой — снижение уровня защищённости частного жизненного пространства.

Совместные сетевые действия значительно усиливают эффект открытия данных. Такой эффект очень незначителен, если это делает лишь одна организация.

Полезность открытых данных НКО — свидетельство зрелости третьего сектора, его способности консолидировать усилия, координировать свою деятельность.

Обобщённо, **модель устойчивого открытого обмена данными между НКО** включает 3 компоненты:

1. Техническая система: наличие онлайн-платформы, которая предоставляет набор простых и надёжных инструментов использования и пополнения данных и обеспечения их сохранности.
2. Достаточно высокая общественная ценность сформированного массива открытых данных.
3. Наличие конкретной заинтересованности участников, от которых зависит наполнение системы данными.

Отсутствие или слабость любой из этих компонент приводит к потере устойчивости системы, а усилия, затрачиваемые на создание и поддержание системы, не оправдываются.

Перспективы развития систем открытых данных просматриваются не очень ясно. НКО не имеют средств для развития в этом направлении. Открытые данные и вообще работа с данными не входят в перечень наиболее приоритетных проблем организационного развития, которые решают НКО.

В развивающихся странах подавляющее большинство НКО решают задачи выживания и поддержания на минимальном уровне своей способности исполнять свою миссию.

В развитых странах действует гораздо больше стимулов для открытия данных в интересах повышения эффективности деятельности НКО. Имеется также достаточно развитая инфраструктура поддержки работы с данными – как коммерческая, так и некоммерческая. Это даёт возможность организациям объединяться и платить относительно небольшие деньги за услуги. О создании собственных систем речи не идёт. Есть добровольческие организации специалистов (такая появилась и в России благодаря Агентству социальной информации — «Теплица социальных технологий»²⁵). Есть фонды, которые могут поддержать именно проекты открытых данных.

Очень много зависит от **доноров**.

Они оказываются сильно заинтересованной стороной, поскольку финансируют однотипные проекты разных организаций. Именно они могли бы стать создателями порталов открытых данных. К ним относятся и операторы президентских грантов, а сам портал мог бы быть создан централизованно государством, поскольку операторы не имеют на это средств и полномочий, нет также надобности создавать несколько отдельных порталов. Российский «Портал некоммерческих организаций»²⁶, например, содержит много данных, но являются ли они открытыми?

Частные доноры могут быть заинтересованы в поддержке инициатив по открытости НКО. Минимально, например, размещение в Интернете данных, полученных в ходе профинансированных проектов, в дополнение к тем суммам, которые необходимы на выполнение самого проекта.

Необходимы и специальные проекты, которые будут демонстрировать ценность открытых данных, привлекая внимание и интерес других заинтересованных сторон.

Полезен также пример самых крупных ведущих организаций, демонстрирующих не только открытость (в смысле прозрачности), но и «высокую» культуру данных.

В конечном счёте, предполагается, что именно информационные технологии в ближайшие годы будут тем инструментом, за счёт которого будут совершенствоваться проектные и организационные технологии, не говоря уже о социальных.

Вместо заключения

Резюме

1. «Открытые данные» — это не просто выкладывание массивов в сети. Это — повышение отдачи конкретного массива. Чтобы этого добиться, нужно это

²⁵ <https://te-st.ru/>

²⁶ <http://www.portal-nko.ru/>

планировать. Это поддаётся оценке — социальная значимость (и его оценка в деньгах). Политика открытости измеряется не количеством опубликованных массивов и объёмами, а общественным эффектом.

2. Значимость того или иного массива определяется сообществом его активных и пассивных пользователей. Чтобы усиливать эффект открытия данных, нужно заниматься развитием соответствующего сообщества. И, наоборот, при желании усилить уже развитое сообщество, нудно подумать, каких данных ему не хватает, сделать эти данные более доступными для этого сообщества. Иными словами, можно идти как со стороны данных, так и сообщества, но эти вопросы связаны.
3. Открытие данных — это проект, направленный на то, чтобы сделать конкретный набор данных работающим на благо общества. И его публикация — это только один из шагов в этом направлении.

Открытые данные: власть для НКО

- Обеспечение прозрачности власти, механизм гражданского контроля.
- Повышение информированности по общественным проблемам (для правильной расстановки приоритетов) и отслеживании своих результатов (использование данных, собираемых властью для оценки деятельности НКО).
- НКО-сообщества должны организовать, чтобы сформулировать запрос на открытие данных власти, чётко сформулировать, какая информация и зачем требуется. Тогда можно ожидать ответных шагов. Иначе придётся ждать, какие данные откроют в очередной раз и насколько они полезны. Но для формулировки такого запроса нужен достаточный уровень самоорганизации сектора.

Открытые данные НКО

- Стимулирует самоорганизацию — за счёт большей информированности о том, что происходит вокруг, чем занимаются другие.
- Повышение эффективности: использование общих массивов; согласованный, скоординированный сбор данных; возможность оценки через сравнение (бенчмаркинг); обмен опытом и знаниями (знать, что и как работает).
- Развитие, инновации: быстро узнавать о новых возможностях.
- Ограничение: открывает один — пользуются все (у НКО нет ресурсов, чтобы заботиться о других, на себя не хватает).
- Необходимость продуманных критериев «работающих» данных.
- Чтобы развивать обмен данными между НКО, нужно развивать сообщества. В неразвитых сообществах, вне зависимости от открытости — это работать не

будет (либо очень слабо, с мало заметным эффектом). Однако в ряде случаев само открытие данных подталкивает развитие сообществ.

Немного про Россию

- Жизнь оцифровывается. Все сектора пытаются использовать новые технологии, чтобы повысить общественный эффект.

В России – пока отставание. Рынка анализа данных не формируется, «культура данных» в некоммерческом секторе развивается медленно, внутренних мотивов для такого развития не видно.

В какой-то момент можно ждать всплеска, но этому должно предшествовать понимание того, что новые технологии дают реальный экономический эффект, экономию или выгоду; либо помогают вести дела, находить решения, которые невозможны без новых технологий. Пока же к данным отношение как к средству иллюстрации своей деятельности. Они не рассматриваются, как источник знаний, понимания – что и как делать, чтобы быстрее и эффективнее решать общественные проблемы.

- Инициатива интересна для рынка, когда имеет значимые экономические эффекты. С одной стороны, это приводит к появлению коммерческих игроков. С другой — сами игроки начинают эффективно подталкивать спрос, формируя потребность. Пока такого перелома не произошло, процесс будет развиваться очень медленно и держаться на «плечах энтузиастов».
- Нужны примеры применения новых технологий и реального выигрыша.
- Развитие массивов данных сильно связано с развитием пользовательских сетей и сетей поставщиков данных. Это дуальная система — массивы и сетевые эффекты. Пользователи — они же поставщики.
- Усилия по развитию открытости данных — это усилия по развитию сообществ.