

Предложение за
мерки в рамките за стратегията
за открито управление на
Република България

Технически концепции и
идеи за отваряне на данни

15.04.2010
Боян Юруков

Съдържание

1	Въведение.....	3
2	Основни технически изисквания.....	4
2.1	Портал за отворени данни.....	4
2.2	Технически изисквания за нови системи.....	4
2.3	Технически изисквания за вече съществуващи системи.....	7
3	Предложения в краткосрочен план.....	8
3.1	Бюджети.....	8
3.2	Агенция по приходите.....	9
3.3	Обществени поръчки и еврофондове.....	9
3.4	Агенция по вписванията.....	10
3.5	Данни на кабинета.....	10
3.6	Здравна инспекция.....	11
3.7	Административен регистър.....	11
3.8	Законова рамка.....	12
3.9	Портал	12
3.10	Незабавни технически поправки в системите.....	12
4	Предложения в дългосрочен план.....	15
4.1	Агенция по вписванията.....	15
4.2	Агенция по приходите.....	15
4.3	Данни от кабинета.....	15
4.4	Министерство на Вътрешните Работи.....	16
4.5	Министерство на здравеопазването.....	16
4.6	Природни ресурси.....	17
4.7	Механизъм за идентифициране на полезни данни.....	17
4.8	Съдействие с чужди организации.....	18
5	Допълнителни връзки.....	20

1 Въведение

Отворените данни са предпоставка за истинска прозрачност в едно управление. Публикуването на стенограми, документация и статистика е само първата стъпка. Предоставянето на необработени, структурирани, документирани и свободни за употреба данни дава възможност журналисти и граждани да смесват, визуализират и анализират информацията, на базата на която се взимат решения в една администрация.

Това е невъзможно при публикуването на статистика и огромно количество документи в свободен текст. Затова отворените данни трябва да са таблици, бази данни, регистри и семантично описани структури. Точно това разграничава откритото управление (*opengov*) от отворените данни (*opendata*) – при първото имаме достъп до информация за работата на държавните структури, докато при второто вече обществото може да проверява решенията и изводите на администрацията, да комбинира данните по неочакван начин и благодарение на това да съветва властимащите и да открива нови бизнес възможности. В научният труд на Харлан Юу и Дейвид Робинсън от Принстън се описва доста добре историята и разграничението между тези две понятия (*Секция 5. – I*).

Отварянето на информацията има голям принос и в работата на самата администрация. Често чиновниците могат трудно да стигнат до качествени данни, на базата на които да взимат решения. Формализирането на процеса, документирането на структурата им и разграничаването на отворени и защитени данни помага за по-доброто управление на достъпа до информация в държавните структури. Това както ще помогне на откриването на полезни данни, така и ще предотврати изтичането на секретни и лични такива.

2 Основни технически изисквания

Това са предпоставки, които ще са нужни за устойчивото разработване на отворени данни в държавната ни администрация.

2.1 Портал за отворени данни

Трябва да изготвим централен портал, който да съдържа връзки към данните отворени в рамките на държавната администрация. Има готови платформи с отворен код, които вече се използват от много държави по света. Препоръчвам платформата SKAN (<http://ckan.org/>) разработена от Фондацията за Отворено Знание (*Open Knowledge Foundation* или *OKFN*). Тази платформа предоставя възможност за лесно документиране и управление на данните. В нея може лесно да се опише структурата, качеството на данните, източника, отговорната институция и ограниченията за използване. Позволява добавянето на данни от неправителствени организации, приложения и визуализации създадени от граждани и примери за успешни проекти използващи такива данни. Порталът ще даде възможност лесно да се намира цялата нужда информация както от страна на работещите в публичния сектор, така и от граждани и журналисти.

Тъй като платформата е достъпна свободно, създаването на такъв портал ще изисква минимални разходи и работа. Дизайнът е налице, но може да се адаптира, за да наподобява повече правителствените сайтове. Преведен е вече на български език и се обогатява постоянно с функции. Съдържа прости инструменти за анализ и визуализация.

2.2 Технически изисквания за нови системи

Държавната и местната администрация работят с голям брой системи и бази данни. Те са изградени през годините с помощта на различни технологии и това често е пречка за интеграцията им. Затова имаме нужда от ясни технически изисквания за всяка нова система. Това ще помогне както за включването им в системите електронното управление, така и за отварянето на част от данните, с които работят.

2.2.1 Технически дефиниции

Първата стъпка е създаването на речник с технически дефиниции свързани с отворените данни. Той ще включва описанието на термини, технологии и стандарти. Трябва да бъдат документираны възможностите, ограниченията и ползите от всяка от тях. На първо място трябва да се даде точно

определение за отворени данни и каква информация може да се съдържа в тях. То може да изглежда така:

Публично достъпна информация с високо качество предоставена за свободно ползване в структуриран и документиран формат базиран на отворени интернет стандарти. Тази информация може да описва както термини, категории и структурите в държавата, така и да предоставя оперативни данни в реално време като бюджети, регистри и събития.

Някои от технологиите, които трябва да се опиташ включват: SQL, XML, CSV, Excel таблици, XML, XSLT, RDF, OWL, WSDL, HTML. Това ще позволи на работещите с тях да имат общо разбиране за терминологията и изискванията към новите системи.

Трябва да се дефинира и какво представлява качеството на данните:

Добре документирана, актуална и точна информация използваща единен речник на термини и категории.

Има нужда и от ръководство какви данни могат да се публикуват, както и как да се изчистват масиви от данни от лична информация. Това може да става чрез групиране на данните в различни периоди от време, по региони и по категории. Така ще се предотврати идентифицирането на частни и юридически лица и в същото време ще запази качеството на данните.

2.2.2 Ръководство за отваряне на данни

Трябва да се изготвят напътствия към изготвящите обществени поръчки и изисквания за нови информационни системи. С тяхна помощ ще може да се идентифицират публично достъпната информация. Това може да става автоматично – създаването на справки от базите данни и публикуването в правителственият `opendata` портал, както и с въвеждането на информационни услуги за достъп до тях в реално време. Режимът за достъп до тези данни може да бъде регламентиран в порталът и да става с регистрация и безплатното получаване на ключ за достъп. Така ще се предотврати претоварването на системи и ще даде възможност за максимално качество и актуалност на получената информация. Пример за данни, на които може да се прави справки, са бюджетна информация, природни ресурси и концесиите върху тях. Примери за данни за реален достъп са търговският регистър, обществените поръчки, болничната информация и данните за заетите лица в бюджетната сфера.

2.2.3 Изготвяне на единна схема за данните в държавната администрация

За да се публикуват данните в отворен формат, има нужда да се изготви обща схема. Пример за такава схема е форматът, в който се предоставя информация за бюджетите, за фирмите в търговският регистър, за кадастъра, за стенограмите и решенията на кабинета и прочие. Общата схема ще улесни както изграждането на нови информационни системи, така и комбинирането и унифицирането на данни между различните институции. Техническите спецификации зависят от форматът на данните – таблици или дървовидни структури. С тях ще се описват колоните и елементите, какво означават, какъв вид данни включват (текст, цифри, дати) и как се групират.

Така всяко ведомство ще публикува информацията си в единен формат, което ще позволи интеграцията и сравнението им. Такъв формат може да се използва и за информационните услуги, където ще се определи единен режим и формат за достъп до всички. Това ще позволи използването им в едни и същи приложения без нуждата от допълнителни технически усилия за интеграция на данните на всяко отделно ведомство.

Важно е да се отбележи, че вече съществуват стандарти и описания използвани в други държави. Поради разликата в работата, структурата, законодателството и дефинициите във всяка от тях, някои от тези схеми трябва да се адаптират за условията в България. За целта е нужен опита на хора запознати с работата на всяко ведомство, както и с конкретните информационни системи. Процесът за дефиниране на единни схеми на данните не е еднократен, а продължителен. Една схема може да се променя и обогатява с времето при промени в законодателството, структурата на администрацията и отварянето на нови данни.

2.2.4 Организация и отговорност

Трябва да се определи структура или човек отговарящ за актуалността и качеството на всички данни във въпросното ведомство. Така ще се децентрализира отговорността и ще може лесно да се проследят проблемите с различни данни. Такава отговорност вече е въведена в регистъра за ражданията в Министерството на здравеопазването където във всяка болница има лице отговорно за въвеждането на данни. Това лице има осигурен достъп до конкретната система и така може лесно да се проследят промените и свършената работа.

Този принцип може да бъде дефиниран формално за всички такива регистри и бази данни и сам по себе си да бъде отворен като данни.

Изготвяйки списък на отворените данни в администрацията, към всяка публикувана таблица освен данни за последно обновление и източник, може да се добави отговорното лице или институция. Така ще се припише отговорност, лице за контакт в случай на проблем и въпроси. Това ще гарантира прозрачност и ангажираност към качеството на данните и използването им от обществото.

2.3 Технически изисквания за вече съществуващи системи

Принципите дефинирани за новите системи могат да бъдат използвани и тези, които вече съществуват. Доста от тях вече предоставят възможност за изготвянето на справки и дори предоставят услуги за достъп в реално време.

2.3.1 Идентифициране на данни за отваряне в съществуващи системи

Трябва да се документират справките и услугите използвани в сегашната администрация. След като се прецени чувствителността на информацията – дали съдържат секретни или лични данни – може директно да се предоставят за ползване в рамките на държавният opendata портал. Такива системи има в повечето регистри, таблици публикувани от статистическият институт, кметства и областни управи. При обновяването на информационните системи трябва да се вземе под внимание тези услуги и справки и да се използват дефинираните отворени стандартни и схеми, за да се осигури унифицирането и надеждността им.

2.3.2 Трансформиране на съществуващите данни

За да се осигури обща структура на данните (единни категории и термини, колони в таблиците) в някои случаи ще е възможно да се променят просто справките, които се изготвят автоматично в тези системи. Пример за това е сайта на парламента, където вече се публикуват данни за работата на комисиите и в пленарна зала. За жалост данните не са документирани, често не са обновени или съдържат грешки. Може да се приеме и адаптира структура използвана в други парламенти и да се осигури единна точка да за достъп до всички данни.

Там където трансформирането на вече съществуващи справки и услуги в самите информационни системи не е възможно, може да се създаде трансформация в рамките на държавния opendata портал. Това става чрез използването на малки програми, които взимат вече изготвените таблици, превеждането и групирането на категории и термини и публикуването на данните във вече дефинираният отворен формат.

3 Предложения в краткосрочен план

Тук са изброени данни, които ще могат да бъдат отворени след минимални организационни или технически промени. Във всички случаи данните са налице и се изисква единствено механизъм за публикуването им в интернет.

3.1 Бюджети

Приходите и разходите в бюджетната сфера бъдат доста тревоги у гражданите и прозрачността ще донесе не само яснота, но и по-голяма подкрепа.

3.1.1 Единен формат

Трябва да се изготви стандартна таблица, в която всяко ведомство да отбелязва бюджета си. Това може да е Excel таблица, в която редовно да се попълват данни. Стандартизирането на този формат и придържането към него ще улесни анализа и събирането на данните. Когато всички числа се попълват на правилните места, таблиците могат лесно да се превърнат в други формати позволяващи по-лесен автоматичен анализ.

3.1.2 Делегирани бюджети

Трябва да се публикува в отворен формат разпределението на държавният бюджет по пера. По аналогичен начин може да се дефинират планираните приходи и източниците им. Обновен доклад за изпълнението на бюджета може да се попълва всеки месец. Това ще помогне да се покажат промените през годините, изпълнението на бюджета през месеците и кои сектори се справят най-добре. Този тип анализ вече се предлага в платформи като <http://openspending.org> и <http://wheredoesmymoneygo.org>. Вторият сайт предлага на посетителите да видят точно колко от данъците им отиват за услуги, пенсии, администрацията и прочие.

3.1.3 Бюджети във всяко ведомство

По аналогичен начин трябва да се публикуват приходите и разходите във всяка институция и агенция. Определени пета и бюджети на сектори са секретни и не могат да се дават подробности. В други има лични данни. Може обаче да се даде обща картина по категории – заплати, бонуси, наеми и канцеларски материали, командировки, поддържане на информационни системи, приходи от лицензи, наеми, такси и дарения.

Обновяването на тези данни ежесечно ще позволи добър контрол както от страна на гражданското общество, така и на кабинета и самите служители. Често тази информация е принципно достъпна, но самото ѝ получаване е трудно и бавно.

3.1.4 Нормативна рамка

Документиране на това какъв е режима за работа с излишъците и приходите. Видяхме, че това е важна тема и съществуват съмнения, че излишъците се харчат съмнително. Внасяне на яснота чрез ясно дефиниране на това как текат паричните потоци в бюджетната сфера би решило този проблем.

3.2 Агенция по приходите

Информационните системи на НАП и скорошното свързване на всички касови апарати са богат източник на полезни данни, като в същото време се запазва анонимността на частни и юридически лица.

Може да се публикува ежесечен доклад с приходите по общини и сфери на дейност – производство, търговия, услуги и прочие. Това ще позволи проследяването както на икономическата активност на гражданите, така и ефективността на мерките на НАП. Възможно е също така да се публикува информация за задълженията, данъчните кредити и върнатото ДДС, за да придобием по-добра представа за измеренията на тези процеси.

Въпреки, че често чуваме цифри и статистика за работата на НАП, често ни е трудно да ги съпоставим и да си ги представим в контекста на бюджета и състоянието на бизнеса. Визуализирайки тези данни ще ни помогне да видим общата картина по-добре.

3.3 Обществени поръчки и еврофондове

И в момента системата за обществени поръчки дава възможност за преглед на обявените и спечелените такива. Порталът за еврофондовете пък предлага добър поглед над перата, по които се водят проекти и има потенциал. Цялата тази информация трябва да бъде представена в отровен формат, за да може да се анализира.

В настоящият момент свалянето ѝ от който и да е от двата сайта е практически невъзможно. Страниците са стандартизирани, което говори, че има база данни зад тях. Ще е много лесно тя да се опише и пусне за свободно сваляне. Вероятно почти цялата информация от нея така или

иначе се показва на сайта – ще е много по-удобно да се предостави направо свободно.

Трябва също така да се пусне информация за спечелилите поръчките. Към този момент това може да се научи най-често от медиите присъствали на срещите и процедурите. Информацията за конкретните оферти и спечелилите трябва да е достъпна, за да се гарантира най-висока степен на прозрачност.

3.4 Агенция по вписванията

Агенцията по вписванията разполага с няколко регистъра, които съдържат изключително ценна информация. На първо време можем да се спрем върху търговският регистър и да го отворим изцяло. Към този момент той е публичен, но практически неизползваем през интернет сайта си. Използването му се блокира както от лошият интерфейс, така и от анти-спам защита, която постоянно трябва да се въвежда.

Да се свали целият регистър струва 30000 лева плюс такса за всяко обновяване. Към този момент 4 (четири) фирми са платили тази цена, а само една получава обновления за всяка фирма. Тези приходи не оправдават дори техническата поддръжка на системите равняваща се на стотици хиляди лева, които заедно с административното поддържане се плащат от бюджета. В същото време, дори държавната администрация, която вече има всички данни, плаща стотици хиляди за софтуер на същите тези фирми и чрез него получава достъп до същите тези данни.

Вместо това, може да се предостави за свободно сваляне цялата база данни съдържаща общата информация за фирмите – име, регистрация, предмет на дейност, собственици, история и прочие. Документите съдържащи лични данни и подписи може все така да са достъпни с електронен подпис. Това ще намали неимоверно натоварването на сървърите. Най-важният ефект обаче е, че ще отвори вратите за по-добри и безплатни интерфейси за достъп. Данните ще могат да се анализират от всеки за скрити схеми, фирми фантоми и връзки между бизнесмени.

3.5 Данни на кабинета

Както всяко ведомство, и кабинета ще може да публикува информация за оперативните си разходи, персонала и обществените си поръчки. Допълнително може да се предостави в отворен формат данните от

системата за управление на документи и решения. По същият начин може да се индексират стенограмите като се отбележи в структуриран формат кой е взел думата и каква тема се обсъжда в дадения момент.

3.6 Здравна инспекция

Към този момент Министерството на здравеопазването поддържа много регистри, а здравната инспекция пази данни за проверките и нарушенията на болници, лични лекари и ресторанти.

Възможно е на първо време да се публикува тази информация за големите болници, както и да се опишат регистрите и да се пускат базите данни, в които се пазят за свободно сваляне. Аналогично на другите примери, те се показват на сайта в структуриран вид, който обаче не позволява свалянето и използването им.

Инспекцията пък публикува редовно доклади за заболяемостта и статистика за нарушенията и проверките на различни заведения. Тези данни се публикуват в графичен вид и отново не могат да се използват. Може лесно да се публикуват Excel таблиците, с които работят.

3.7 Административен регистър

Сайта на административният регистър позволява да се следи в реално време заетите в държавната администрация, както и свободните места. Невъзможно е обаче да се получат пълните данни за всички министерства, областни управи или кметства. Интерфейса е труден за използване и може да се вижда само информация за една структура на веднъж.

Затова трябва да има възможност да се свали цялата база данни, заличавайки имената на конкретно заетите на въпросните длъжности. Разбира се, тези на ръководни постове ще останат в данните. В крайна сметка ще получим същата информация, но удобна за работа прескачайки представянето ѝ на самият сайт. Тази база данни съществува и е въпрос на една заповед да бъде предоставена свободно за сваляне.

За да се намали натоварването на системите при по-често сваляне, възможно е да се изготвят само справки за промените в последната седмица. Така когато някой свали базата данни веднъж, ще може просто да сваля редовно промените. Така ще се получи историческа картина на промените в държавната администрация – нещо несъществуващо към този момент.

3.8 Законова рамка

Предоставянето на отворени данни от страна на държавните институции трябва да се регламентира в законодателството. Сега служителите се притесняват да предоставят дори данни, които са публични и би трябвало да бъдат достъпни.

Това може да стане чрез въвеждането на възможност за изискване на пълна база данни или редактирана част от нея по Закона за достъп до публичната информация. Може също така да се въведе изискване, че когато поискана справка по този закон не е свойствена за ведомството, то тогава вместо нея да бъде предоставяна всички данни нужни за изготвянето ѝ.

3.9 Портал

Създаването на държавен портал за отворени данни е от особена важност. Той ще позволи намирането на всякакви данни както от журналисти и граждани занимаващи се с анализ на данни, така и от служители в администрацията търсещи източници на информация. Подходяща платформа за целта е SKAN, която споменах по-горе. Тя вече се използва от редица държави за техните opendata сайтове. Българският може да е например opendata.gov.bg.

Този портал може веднага да се запълни с отворени данни като тези на парламента, резултатите от изборите в ЦИК, таблиците на НСИ и на финансовото министерство. Ще е добре да се добави и секция за данни отворени от НПО-та и частни лица, както и галерия с връзки към полезни визуализации и инструменти за анализ. Такива лесно могат да се съберат с помощта на посетителите на сайта, както и партньори като OKFN и Sunlight Foundation.

3.10 Незабавни технически поправки в системите

В не малко сайтове на държавната администрация има проблеми, които могат лесно и бързо да бъдат решени.

3.10.1 Без анти-спам

Доказано е, че captcha или анти-спам проверките не работят. Сравнително лесно могат да се прескочат или измамат. В същото време те блокират работата на мнозинството от посетителите. Съвременните интернет сървъри имат достатъчно добър контрол над разпределение на натоварването и блокирането при съмнения за атака.

Един отрицателен пример в тази насока е регистъра за ражданията на Здравното министерство. В него се пазят данни за всички раждания и се обновява в реално време. На сайта се публикуват пак в реално време цифри обобщени по области. От една страна, сайта не може да се отвори в чужбина. По-важното е, че до скоро не данните не можеха да се свалят. След като открихме начин да ги свалим без да натоварваме сървъра им, направихме по-добра и интерактивна графика сравнявайки ги с населението в определена възраст и раждаемостта в минали години: <http://opendata.yurukov.net/birth/>

За жалост обаче сайта ни беше първо блокиран от страна на администраторите, а после беше въведена анти-спам защита. Така сега трябва да прекарваме по няколко часа седмично да свалим същите данни, които иначе би трябвало да са свободно достъпни. Като резултат, сайта не е обновяван от два месеца.

Точно това е единственият ефект на анти-спам проверките и затова трябва и могат да се свалят бързо.

3.10.2 Кратко описание за принципите на отворените данни

Администраторите на информационните системи в различните ведомства трябва да бъдат запознати с принципите на откритото управление. Точно те ще са тези, които в крайна сметка ще отворят данните. Затова ще е добре от рано да се информират и насърчат да търсят информация. Често самите те биха предложили данни за отваряне и биха търсили допълнителна информация в портала. Това ще помогне на по-късен етап и в идентифицирането на ключовите субекти в държавната администрация за описването на единни дефиниции и схеми за данните.

3.10.3 Указание за пълно сваляне на данните

Както споменах в няколко точки по-горе, съществуват редица бази данни, които могат направо да бъдат пуснати свободно в мрежата. Често те се поддържат в периферията на информационните системи и в тях се съхраняват единствено справки без оперативна информация или лични данни.

Макар интуицията да диктува, че тези бази данни съществуват, за да представят данните в текстови вид чрез интернет страници, това далеч не е достатъчно. Самите бази данни са ресурс и запазването на оригиналната структура е нещо ценно.

Затова може да се издаде указание за това как тези бази данни да бъдат изчиствани, извеждани и предоставяни в публичното пространство. Това може да става и автоматично, тъй като доста от системите за бази данни имат такава функция. В указанието може също така да се определи в кои случаи да се предоставят в съкратен вариант обновленията и в какъв формат.

3.10.4 Прекратяване на публикуването на PDF документи

Когато става дума за текстови документ, използването на PDF като универсален формат за четене е препоръчителен. Често обаче той се използва и за публикуването на таблици с данни. Това ги прави практически безполезни.

Пример за това е парламента, който от миналата година започна да публикува данните за гласуването на депутатите. В началото те бяха под формата на таблица в PDF документ, но след молба от наша страна, администраторите започнаха да публикуват и Excel таблицата, от която така или иначе се създаваше PDF документа,

3.10.5 Поправяне на лицензите на правителствените сайтове

Все още много от интернет сайтовете на правителството, министерствата, областните управи и агенциите декларират запазени права над съдържанието и данните си. Това означава, че – де юре – те не могат да бъдат копирани и използвани по никакъв начин. Докато това най-вероятно е пропусака, то предразполага към недоразумения. По смисъла на закона всичко публикувано на тези страници е собственост на обществото и това трябва да бъде описано като такова на самите страници.

Интересното е, че дори страниците на различните съдилища посочват запазени права над сайтовете си, а някои от тях са създадени на базата на платформи с отворен код. Това е нарушение както на законите на Република България, така и на свободните лицензи на самите платформи.

4 Предложения в дългосрочен план

Това са идеи, реализацията на които ще изисква време, промени в регламенти и нормативни актове или организация на събития.

4.1 Агенция по вписванията

Всички регистри на агенцията по вписванията могат да бъдат предоставени свободно – в пълния или съкратения си вид. Дори когато вече има платен режим за достъп, съществува част от данните, които е достъпна, но трудна за достъп. Пример за това е Имотния регистър, където част от данните могат да се търсят, но не и да се свалят. В интерес на прозрачността, в публично достъпните данни трябва да се включи и информация за физическите и юридическите лица собственици на имотите.

4.2 Агенция по приходите

Публикуването на информация за приходите и задълженията е първата стъпка. Свързването на всички касови апарати предоставя златна мина от данни. Възможно е да се групират данните по общини и седмици така, че да не може да се идентифицират отделни търговци и фирми. Ако има начин да се категоризират апаратите към определен регион или сектор от бизнеса, то може да се изготвят подробни ежеседмични доклади за приходите от търговска дейност с граждани. Това ще е безценен източник на информация за реалната покупателна способност на българина не само като цяло, но и като разпределение по места и във времето.

4.3 Данни от кабинета

Възможно е да се изготви категоризация на целите и приоритетите. Така ще се опишат заявените задачи, стъпките им и напредъка по изпълнението им. Такива проекти вече има в интернет пространството. Графичното проследяване на заявените цели ще помогне за по-добра прозрачност, поставяне на конкретни решения в контекст, категоризиране на документацията и по-добра видимост за дългосрочният план на действие.

Всичко това може да бъде описано в отворен формат, визуализирано с графика и предоставено на други за анализ и потвърждение.

Допълнително, може да се изготви индекс на отвореност на институциите. Този индекс ще съдържа критерии за това каква част от публичните данни да в отворен формат, какво качество на данните съобщават използващите

го и до колко ръководството е ангажирано с прозрачността и откритото управление. Така ще може да се оценяват агенции, министерства, областни управи, кметства и прочие и да се насърчи работата им в тази насока.

4.4 Министерство на Вътрешните Работи

Полицията работи с много оперативни данни, които не могат да бъдат публикувани. Определянето обаче на тези, които са публично достояние и разграничаването им от останалите може дори да подобри работата със секретни данни.

Всяко управление изготвя справки за работата си, но често те са лишени от контекст и остават незабелязани. Паралелно с това се публикуват ежедневно бюлетини в свободен текст пръснати из множество сайтове, които описват престъпления и инциденти станали през деня. Всичко това не позволява никаква работа с данни на министерството.

За целта може да се изкарва ежедневна справка в отворен структуриран формат за всички престъпления, нарушения по пътищата и издирвани лица. Включвайки категории като тип престъпление, дали има смъртни случаи при катастрофа, дали е използван алкохол или колан, както и включване на относително географско местоположение на случая ще помогне много. Този тип данни се използват отдавна в щатите за анализ както на престъпленията по региони, така и за проблемите на пътищата. Разглеждайки всеки отделен доклад е трудно да се види общата картина и да се вземе информирано решение.

В същото време ще е възможно обществеността да потвърждава и надгражда над изводите на полицията. Ще се прецени ефективността на определени мерки за борба с престъпността и катастрофите, ще се повиши информираността у хората и ще се съберат предложения. Така ще се използва на обществото от доброволци и анализатори за неща, които полицията може би няма капацитет.

4.5 Министерство на здравеопазването

Трябва да се публикуват данни за всички проверки, нарушения, предписания и време за решаване на всички болнични и заведения за хранене. Това от една страна ще има дисциплиниращ ефект, а от друга ще даде ценна информация на потребителите. В Монреал един сайт е направил така карта с проверените ресторанти позволяваща на всеки да провери дали има риск и да реши дали да яде на дадено място.

Паралелно с това може да се публикуват отворени данни за лекарствената листа, сигналите за корупцията, клиничните пътеки, плащанията към болници и лекари, приходите на Здравната каса и прочие.

С въвеждането на електронната здравна карта, прегледите, а потенциално и клиничните пътеки на всеки болен ще се получават централно. При максимална сигурност на личните данни може да се изградят справки за общата заболяемост по общини и дори квартали ежедневно или ежеседмично. Така ще може наистина в реално време да се засичат както епидеми, така и по-дребни проблеми. Ще може ясно да се видят и повишаването на болничните около празници и отпуски.

4.6 Природни ресурси

Отворените данни може да бъдат статични и динамично. Статичните представляват дефиниции, класификация и географски данни, които се променят изключително рядко. Такива може да са данните за природните ресурси. На първо място трябва да се документира в отворен формат видовете, мащабите, свързаните рискове и режимите на концесия на различните природни ресурси. След това може да се документира местоположението, дълбочината и мащабите им. Накрая след като се изгради регистъра ще може да се предостави информация за концесионерите, предишните такива и наказанията върху тях.

4.7 Механизъм за идентифициране на полезни данни

Ценното в отворените данни е, че информацията получена от тях може да бъде използвана по неочаквани начини. Това е силата на истински откритото управление. В този смисъл обаче може да съществуват данни в администрацията, които остават неизвестни, но може да представляват интерес за обществото. Пример за това е Административният регистър, който е почти непознат, но съдържа много ценна информация.

За целта трябва да се изгради продължителен механизъм в рамките на ангажимента за отворени данни, с който да се приемат предложения за отваряне на данни както от страна на обществото, така и от служители на самата администрация. Технически това може да става през порталът strategy.bg или през порталът за отворени данни.

В рамките на този процес, може да бъдат идентифицирани специалисти вътре в държавната администрация, които имат нужните знания и опит да

боравят с информационните системи и да публикуват структурирани данни. Липсата на такива лица в конкретни ведомства ще говори за недостиг на специалисти поддържащи системите, тъй като уменията често са близки.

4.8 Съдействие с чужди организации

Черпенето на опит от местни или чуждестранни държави е важно също толкова, колкото споделянето на данните на нашата администрация.

4.8.1 Съвместни проекти с НПО-та

Независимо дали са български или чуждестранни организации, сътрудничеството по различни проекти в сферата на откритото управление е важно. Много често тези организации имат много по-голям опит с материята от администрацията. Тези проекти може да включват обучения, конференции, технически проекти и прочие.

При работата с международни организации често се споделят проекти и данни. Съществуват отворени платформи, които могат да бъдат използвани, портали за данни, които събират огромно количество данни от цял свят и идеи, които лесно могат да бъдат трансформирани за нашата среда. Интегрирането на нашите данни с тези от други страни ще позволи неочаквани анализи и изводи. В същото време идеите, контактите и нуждата от още по-добри инструменти и графики ще отвори ниша в пазара у нас.

4.8.2 Обучения на държавни служители

Техническата грамотност е важна за работата със структурирани данни, но също толкова е важно да се разбере смисъла от тях. От една страна трябва да се обучат хората поддържащи и изграждащи изисквания за информационни системи как да дефинират и публикуват отворени данни. От друга страна, служители на администрацията трябва да бъдат научени къде да ги намират, как да използват прости инструменти и как да извличат полза от тези данни за работата си.

4.8.3 Обучения и насърчаване извън администрацията

Използването на данните в администрацията е също толкова важно, както и извън нея. Може да се направят работни групи, обучения и състезания за разследваща журналистика, анализ, визуализация, интернет сайт или софтуерно решение използващо данните публикувани от правителството.

Данните са изключително полезни, но сами по себе си са трудни за четене, макар и в добре структуриран вид. Затова е важно да се стимулира графичното изобразяване на аспекти от тези данни, свързването им с други такива и създаването на интерактивни приложения, които всеки би могъл да разбере и използва. Това ще осмисли ползата от публикуваните данни и ще открийте вратата към крайната степен на прозрачност и сътрудничество между гражданското общество и администрацията.

4.8.4 Организирането на конференции и семинари

В международен план се правят доста конференции за различни аспекти на отворените данни – правителствени, географски, научни и прочие. Те позволяват не само да се популяризират усилията на държавата-домакин в тази насока, но и да се споделят опит, иновации и специалисти. Не малко решаващи стъпки, идеи и приложения са произлезли точно от такива контакти и конференции. Организирайки такива събития ще ни позволи да привлечем опит и специалисти, които да ни помогнат с напредъка по откритото управление.

5 Допълнителни връзки

- I. The new ambiguity of "Open government" - Harlan Yu, David Robinson – Princeton University Feb 2012 - http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2012489
- II. <http://openspending.org/> – портал за бюджетни данни от цял свят
- III. <http://wheredoesmymoneygo.org> – британски сайт за осмисляне на бюджетните разходи в контекста на личният принос на всеки гражданин
- IV. <http://ckan.org/> - платформа с отворен код за портал-регистър за отворени данни
- V. <http://okfn.org/> - международна фондация фокусирана върху отворените данни. Тя разработва инструменти, изгражда контакти, организира събития и събира отворени данни по цял свят.
- VI. <http://opendatahandbook.org/> - Наръчник за отворени данни описващ основните термини, принципи и ползи от публикуването и работата с тях
- VII. Guardian Data blog - <http://www.guardian.co.uk/data> – пример за журналистика на данните разкриваща интересни факти произтичащи от отворените данни на британските власти