



РЕПУБЛИКА БЪЛГАРИЯ
ДЪРЖАВНА АГЕНЦИЯ „ЕЛЕКТРОННО УПРАВЛЕНИЕ“

Утвърдени

със Заповед на председателя
на Държавна агенция „Електронно управление“

№ ДАЕУ-2073/10.02.2020 г.

МЕТОДИЧЕСКИ УКАЗАНИЯ

за дефиниране на структурата и съдържанието
на наборите от данни за публикуване
в Портала за отворени данни
и тяхното актуализиране

София, 2020

СЪДЪРЖАНИЕ

1. Речник на термини, дефиниции и съкращения	3
1.1. Използвани акроними	3
1.2. Технологични дефиниции.....	4
2. Въведение.....	7
3. Политика в областта на информацията от общественния сектор в машинночетим отворен формат (Политика за отворени данни)	8
4. Основни принципи за производство и разпространение на данни от публичната администрация.....	9
4.1. Отворени по замисъл и по подразбиране (Open by design and by default).....	9
4.2. Откриваеми (Findable).....	9
4.3. Достъпни (Accessible).....	10
4.4. Семантично оперативно съвместими (Interoperable).	10
4.5. Повторно използваеми (Reusable).....	10
5. Подготовка и публикуване на набори от данни в Портала за отворени данни. ...	11
5.1. Процес за вземане на решение за публикуване на отворени данни в Портала за отворени данни.....	11
5.2. Оценка на отворените данни.	13
5.3. Стандартни условия за повторно използване на информацията от общественния сектор и за нейното публикуване в отворен формат (стандартни лицензи).....	13
5.4. Препоръчвани формати за отворени данни.....	14
5.5. Използване на схема за метаданни в областта на отворените данни.....	16
5.6. Използване на стандарти за отворени данни.....	18
5.7. Уникални адреси за идентификация на ресурсите (Uniform Resource Identifiers).....	19
6. Актуализиране, поддържане и достъпност на данните и ограничения за публикуване.....	19
6.1. Задължително поддържащо актуализиране на публикуваните данни.....	19
6.2. Създаване на постоянен, устойчив достъп до данните.	19
6.3. Ограничения при разпространението на публичните данни като отворени.	19
7. Практически насоки за вътрешни правила в организациите от общественния сектор с цел организиране на дейностите в областта на отворените данни.	21
7.1. Приложение на принципите за отворените данни.....	21
7.2. Актуализиране на данните.....	22
7.3. Споделяне на данни чрез веб услуги.....	22
7.4. Определяне на длъжностно лице или лица, отговорни за политиката за отворени данни в рамките на организацията от общественния сектор.	22
8. Начини за публикуване на информация на Портала за отворени данни.	23

1. Речник на термини, дефиниции и съкращения

1.1. Използвани акроними

Акроним	Описание
ЗЕУ	Закон за електронното управление
ЗДОИ	Закон за достъп до обществена информация
НОИИСРЕАУ	Наредба за общите изисквания към информационните системи, регистрите и електронните административни услуги
ПОД	Портал за отворени данни
ЕС	Европейски съюз
ЕК	Европейската комисия
W3C	World Wide Web Consortium
ISO	International Organization for Standardization
IETF	Internet Engineering Task Force
OGC	Open Geospatial Consortium
OASIS	Organization for the Advancement of Structured Information Standards
URL	Uniform Resource Locator
URI	Uniform Resource Identifier
API	Application Programming Interface/Приложно-програмен интерфейс

1.2. Технологични дефиниции

Термин	Описание
Уникален адрес за идентификация на ресурс (URI)	Уникален идентификатор, представляващ последователност от символи, идентифициращ абстрактен или физически ресурс.
Софтуер с отворен код	<p>Компютърна програма, която се разпространява при условия, които осигуряват безплатен достъп до програмния код и позволява:</p> <ul style="list-style-type: none"> - използването на програмата и производните на нея компютърни програми, без ограничения в целта; - промени в програмния код и адаптирането на компютърната програма за нуждите на нейните ползватели; - разпространението на производните компютърни програми при същите условия. <p>Списък на стандартни лицензионни споразумения, които предоставят тези възможности, може да бъде намерен в подзаконовата нормативна уредба към ЗЕУ или на: http://opensource.org/licenses.</p>
Машинночетим формат	Формат на данни, който е структуриран по начин, по който, без да се преобразува в друг формат, позволява софтуерни приложения да идентифицират, разпознават и извличат специфични данни, включително отделни факти и тяхната вътрешна структура.
Отворен формат	Означава формат на данни, който не налага употребата на специфична платформа или специфичен софтуер за повторната употреба на съдържанието и е предоставен на обществеността без ограничения, които биха възпрепятствали повторното използване на информацията.
Лиценз	„Лиценз“ се използва като синоним на „условия за ползване“
Стандартен лиценз	Стандартни условия за повторно използване на информацията от обществения сектор и за нейното публикуване в отворен формат.
Информация от обществения сектор	<p>Терминът „информация от обществения сектор“ е синоним на „документ“, „информация“ и „данни“ и представлява всяка информация, обективизирана върху материален (хартиен, електронен или друг) носител, включително съхранена като документ, звукозапис или видеозапис и събрана или създадена от организация от обществения сектор.</p> <p>Термин, аналогичен на термина „документ“ - всяко съдържание, или част от такова съдържание, независимо от носителя му (хартиен носител или електронна форма, или като звукозапис, видеозапис, или аудио-визуален запис), съгласно Директива (ЕС) 2019/1024.</p>

Отворени данни	Това е информация от общественения сектор в отворен машинночетим формат, заедно с нейните метаданни. Концептуално под „отворени данни“ се разбират най-общо данните в отворен формат, които могат да се използват свободно, повторно и да се споделят от всеки за всякакви цели. Тези данни са достъпни безплатно за всеки, добре описани и структурирани са, включително съдържат метаданни и са лесни за машинна обработка.
Набор от данни	Съвкупност от информация, която е съставена от различни елементи, но може да бъде обработвана като една единица от компютър. Наборите от данни могат да съдържат неограничен брой файлове (ресурси) с еднаква структура и предназначение.
Набор от данни с висока стойност	Информация (документи), чието повторно използване се свързва със значителни ползи за обществото, околната среда и икономиката, по-специално поради тяхната пригодност за създаване на услуги с добавена стойност, приложения и нови, висококачествени и достойни работни места, както и поради броя на потенциалните ползватели на услуги с добавена стойност и на приложенията, създадени въз основа на тези набори от данни.
Информационен масив	Група набори от данни с еднакво предназначение или всички данни в една информационна система
Информационен масив за изтегляне	„Bulk download“ представлява група набори от данни, която може да бъде придобита чрез изтегляне в цялост, а не да се предоставя достъп само до разрези от нея.
Ресурс (в ПОД)	Файл с данни с определена структура и предназначение, съставна част на набор от данни.
Метаданни	Данни, описващи структурата на информацията, предмет на повторно използване.
Приложно-програмен интерфейс	Приложно-програмният интерфейс (на английски: Application Programming Interface - API) е автоматизиран софтуерен слой или компонент, с който една информационна система предоставя онлайн достъп до своя функционалност, включително приемане или предаване на данни, на други информационни системи или софтуерни приложения.
Първичен регистър	Регистър, който се поддържа от първичен администратор на данни - административен орган, който по силата на закон събира или създава данни за субекти (граждани или организации) или за обекти (движими и недвижими) за първи път и изменя или заличава тези данни. Например Търговският регистър е първичен регистър за юридическите лица със стопанска цел, Имотният регистър е първичен регистър за недвижима собственост.
CSV	Comma-separated values е файлов формат, подходящ за обмен на големи набори от данни с фиксирана структура.

RDF	Resource Description Framework е стандартен семантичен формат за обмен на данни, дефиниран в група препоръки на W3C.
XML	Extensible Markup Language - разширяем маркиращ език, е стандарт (метаезик), който определя набор от правила за кодиране на документи във формат, който е разбираем за човека и в същото време машинночетим.
JSON	JavaScript Object Notation е файлов формат, който е лесен за прочитане чрез стандартни методи, налични в повечето съвременни езици за програмиране.

2. Въведение.

Този документ се издава на основание чл. 7в, т. 17 от ЗЕУ и представлява методически указания за дефиниране на структурата и съдържанието на наборите от данни за публикуване в Портала за отворени данни и тяхното актуализиране. Той съдържа следните основни раздели:

- Политика за отворени данни;
- Основни принципи за производство и разпространение на данни от публичната администрация;
- Подготовка и публикуване на набори от данни;
- Канали и начини за публикуване на информация;
- Актуализиране, поддържане и достъпност на публикуваните данни;
- Практически насоки за вътрешните правила в организациите от общественния сектор с цел организиране на дейностите в областта на отворените данни.

Основната цел на указанията е да се подпомогнат администрации и организации при структурирането и актуализирането на създаваните от тях набори от данни за публикуването им на Портала за отворени данни съгласно задълженията в ЗДОИ и други специализирани нормативни актове.

Методическите указания се отнасят за организациите от общественния сектор, физическите и юридическите лица, които са задължени субекти по чл. 3 от ЗДОИ, а именно за държавните органи, техните териториални звена и органите на местното самоуправление в Република България, както и за публичноправните субекти, в т.ч. за публичноправните организации и за физически и юридически лица по отношение на извършваната от тях дейност, финансирана със средства от консолидирания държавен бюджет и средства от фондове на Европейския съюз или предоставени от Европейския съюз по проекти и програми.

Целевата група на методическите указания са упълномощените представители на различните организации от общественния сектор и на лицата, отговорни за публикуване на набори от данни в отворен формат на Портала за отворени данни, служителите на задължените организации, създаващи и поддържащи информацията от общественния сектор и техническите специалисти, които организират работата по прегледа, обработката и верификацията на наборите от данни.

Допълнително методическите указания дават насоки на всички организации и лица на територията на Република България, притежаващи права на собственост и/или информация, защитена с авторски права, желаещи да публикуват по собствена инициатива набори от данни в отворен машинночетим формат на Портала за отворени данни в обособен обществен или граждански профил.

Този документ ще бъде от интерес и за онези, които се нуждаят от общо разбиране на процедурата за дефинирането на структурата и съдържанието на наборите от данни за публикуване в Портала за отворени данни съгласно ЗДОИ.

Отворените данни, т.е. информацията от общественния сектор в отворен машинночетим формат, заедно с нейните метаданни, са стратегически ресурс с огромен потенциал за публичния сектор, частния бизнес, академичната общност, неправителствения сектор и гражданите. Те представляват публичната информация, събирана, създавана и поддържана от организациите от общественния сектор и публикувана в отворен машинночетим формат (CSV, RDF, XML, JSON и др.),

позволяващ повторната ѝ употреба. Данните се предоставят с права за свободно използване, включително повторно, обработка, съпоставяне с други отворени данни и разпространение без ограничения. Информацията, която съдържа лични данни, класифицирана информация или друга информация, за която е необходимо да се доказва правен интерес, се изключва.

Национален портал за отворени данни

Националният портал за отворени данни <https://data.egov.bg/> представлява единна, централизирана, публична уеб базирана информационна система, която осигурява публикуването и управлението на информация за повторно използване в отворен, машинночетим формат, заедно със съответните метаданни. Платформата е изградена по начин, който позволява цялостно или частично извличане на публикуваната информация. Данните се разпространяват с лиценз, който позволява свободен достъп, използване за търговски и нетърговски цели и реализиране на приложения.

Развитието на Портала за отворени данни е насочено към публикуване на качествени информационни масиви, набори от данни и ресурси, както и предоставяне на възможности за визуализации и инструменти, улесняващи обработката, публикуването и използването на данните.

3. Политика в областта на информацията от обществения сектор в машинночетим отворен формат (Политика за отворени данни)

Директива 2003/98/ЕО относно повторната употреба на информацията в обществения сектор, изменена с Директива 2013/37/ЕС на Европейския парламент и на Съвета от 26 юни 2013 година, регулира политиката относно повторната употреба на информацията в обществения сектор. Тя е транспонирана в българската нормативна уредба чрез ЗДОИ, където е регламентиран и националният портал за отворени данни като централно хранилище за отворените данни на администрацията. В същия закон, съгласно чл. 15б и чл. 15г, на държавните органи се вменява задължението за приоритизиране и публикуване на данни, предназначени за повторно използване.

На 26 юни 2019 година беше публикувана актуализация на директивата за информацията в обществения сектор, преработена като Директива (ЕС) 2019/1024 на Европейския парламент и на Съвета относно отворените данни и повторното използване на информацията от обществения сектор, т.нар. директива за отворените данни. Новата директива прави редица важни промени:

- Включва данните, произвеждани от публично контролирани дружества в комуналния и транспортния сектори, в обхвата на директивата;
- Разширява обхвата на данните от научните изследвания;
- Ограничава предоставянето на изключителни права на частния сектор върху създадени данни по време на изпълнение на обществени задачи;
- Увеличава прозрачността при предоставяне на изключителни права върху създадени данни по време на изпълнение на обществени задачи.

Една от най-съществените промени е в член 14 относно наборите от данни с висока стойност. Той дава на ЕК правомощия да приеме актове за изпълнение, с които определя списък на конкретни видове набори от данни с висока стойност, които държавите членки ще бъдат задължени да публикуват при отворени лицензи, а в някои случаи и при използване на определени API и стандарти. Актовете за изпълнение могат определят тези стандарти.

Към момента всяка организация от обществения сектор ежегодно планира поетапното публикуване на своите данни на Портала за отворени данни <https://data.egov.bg/>.

Редът и начинът за създаване, поддържане и публикуване на обществена информация в отворен формат и условията за нейното повторно използване са определени в Наредба за стандартните условия за повторно използване на информация от обществения сектор и за нейното публикуване в отворен формат.

Успехът на политиката за отворени данни зависи от три основни фактора:

- качеството на публикуването на отворените данни (напр. точност, пълнота, навременност и последователност);
- използването на отворените данни;
- документиране на въздействието и ползите от отворените данни.

4. Основни принципи за производство и разпространение на данни от публичната администрация.

4.1. Отворени по замисъл и по подразбиране (Open by design and by default).

Всички данни, произвеждани и създавани от или за публичната администрация следва да се разпространяват като отворени по замисъл и по подразбиране. С нормативната база на Република България (като напр. НОИИСРЕАУ) се гарантира, че новите и/или модифицираните информационни системи и регистри, в които се произвеждат и създават данни за или от административните органи, са тяхна собственост и че данните се предоставят/разпространяват в отворен машинночетим формат, освен ако не се предвижда друго в специален нормативен акт или ведомствени нормативни ограничения.

Данните, произвеждани от/за публичния сектор трябва да са отворени и достъпни за обществото за търсене, достъп и използване под отворен, неограничен лиценз, освен ако няма конкретна, непреодолима причина, поради която тези данни не могат да бъдат отворени и тази причина е ясно съобщена на обществеността.

За да се насърчи, популяризира и осъществи политиката за отворените данни в публичния сектор, се препоръчва да се въведе презумпцията за отвореност на данните за повторна употреба, когато няма правни разпоредби, възпрепятстващи това.

В същото време следва да се гарантира постоянно ниво на защита на целите от обществен интерес, като обществената сигурност, включително когато става въпрос за чувствителна информация, свързана със защитата на критичните инфраструктури. Организацията следва също да гарантират защитата на личните данни, включително когато информацията в отделен набор от данни не представлява риск за идентифицирането или посочването на физическо лице, но когато тази информация е в съчетание с друга налична информация, тя би могла да доведе до подобен риск.

4.2. Откриваеми (Findable).

Отворените данни трябва да притежават уникални адреси за идентификация – URI (Uniform Resource Identifier). URI е последователност от символи, идентифицираща абстрактен или физически ресурс. Всеки ресурс и набор от данни в Портала за отворени

данни има уникален идентификатор, автоматично генериран от портала, ако такъв не бъде посочен от организациите при публикуване на набора/ресурса.

4.3. Достъпни (Accessible).

Отворените данни трябва да се намират лесно, включително чрез онлайн търсачките. Отворените данни следва да бъдат достъпни и без бюрократични бариери. Например потребителите не трябва да подават заявка и да чакат индивидуално одобрение от упълномощени служители, за да получат достъп до данните.

Публикуването на данни на централен/национален портал е добра европейска и световна практика, при която се осигурява достъп до всички налични отворени масиви от данни от една единна входна точка, чрез което се улеснява откриването и използването им. В допълнение, публикуването на данни на национален портал гарантира, че данните са достъпни за най-широк кръг от потребители за търсене, достъп, извличане и преглед и последващо повторно използване. В много случаи това включва предоставяне на данните в множество стандартизирани формати, което улеснява обработването им от хора и машини.

В същото време за да се получи лесен и бърз достъп до данните трябва да бъдат разработвани и поддържани приложно-програмни интерфейси. API е набор от функции, процедури, определения и протоколи за междумашинна комуникация и безпрепятствен обмен на данни. За API следва да бъде осигурена ясна техническа документация, която да е изчерпателна и налична онлайн. Когато е възможно, следва да се използват отворени API. Повторното използване и споделянето на данни с помощта на подходящо използвани приложно-програмни интерфейси носи цялостна полза, тъй като помага на разработчиците и стартиращите предприятия да създават нови услуги и продукти.

4.4. Семантично оперативно съвместими (Interoperable).

Семантичната оперативна съвместимост гарантира запазването и разбирането на точния формат и точното значение на данните и информацията при техния обмен. Тя обхваща както семантични, така и синтактични аспекти:

- семантичният аспект се отнася до значението на отделните данни и взаимоотношението между тях. Той включва разработването на речници и схеми за описване на обмена на данни и гарантира, че отделните данни се разбират по еднакъв начин от всички страни;
- синтактичният аспект се отнася до описването на точния формат на подлежащата на обмен информация по отношение на граматика и формат.

Споразуменията за референтни данни под формата на таксономии, контролни речници, тезауруси, списъци с кодове и структури/модели на данни, пригодни за повторно използване, са ключови предпоставки за постигане на семантична оперативна съвместимост. Принципът „отворени по замисъл“, съчетан с технологии за свързани данни, са иновативни начини за значително подобряване на семантичната оперативна съвместимост.

4.5. Повторно използваеми (Reusable)

Всички отворени данни трябва да са съпроводени от информация за произход и лиценз за използване.

Отворените данни, произведени от и за публичната администрация, трябва да се разпространяват безплатно, придружавани от отворен и неограничен лиценз.

Лицензите трябва да бъдат публикувани и свързани със съответните отворени данни, за да се гарантира, че потребителите могат лесно да намират и разбират условията за достъп и повторна употреба на тези данни.

5. Подготовка и публикуване на набори от данни в Портала за отворени данни.

Ключов приоритет за държавата е разработването и развитието на национален портал за отворени данни – „инструмент“, чрез който да се публикува информация от обществения сектор по начин, който я прави откриваема, достъпна, съвместима и подходяща за последващо и повторно използване. Основната цел на портала е да подпомага публикуването на данни с високо качество и обществена стойност, които да задоволяват търсенето и да допринасят за реални ползи за гражданите, бизнеса и публичния сектор в икономически, социален и демографски аспект.

Този раздел описва техническата рамка, върху която се изгражда и развива *политиката за отворени данни* и гарантира, че публикуването на информационни масиви, набори от данни и ресурси на портала <https://data.egov.bg/> се извършва по последователен, непроменлив и прозрачен начин. Документът подлежи на актуализация при всяка промяна на обстоятелствата, вписани в него и/или промени, настъпили спрямо изискванията за прилагане на европейското и националното законодателство, заедно с развитието на технологиите и практиките в областта.

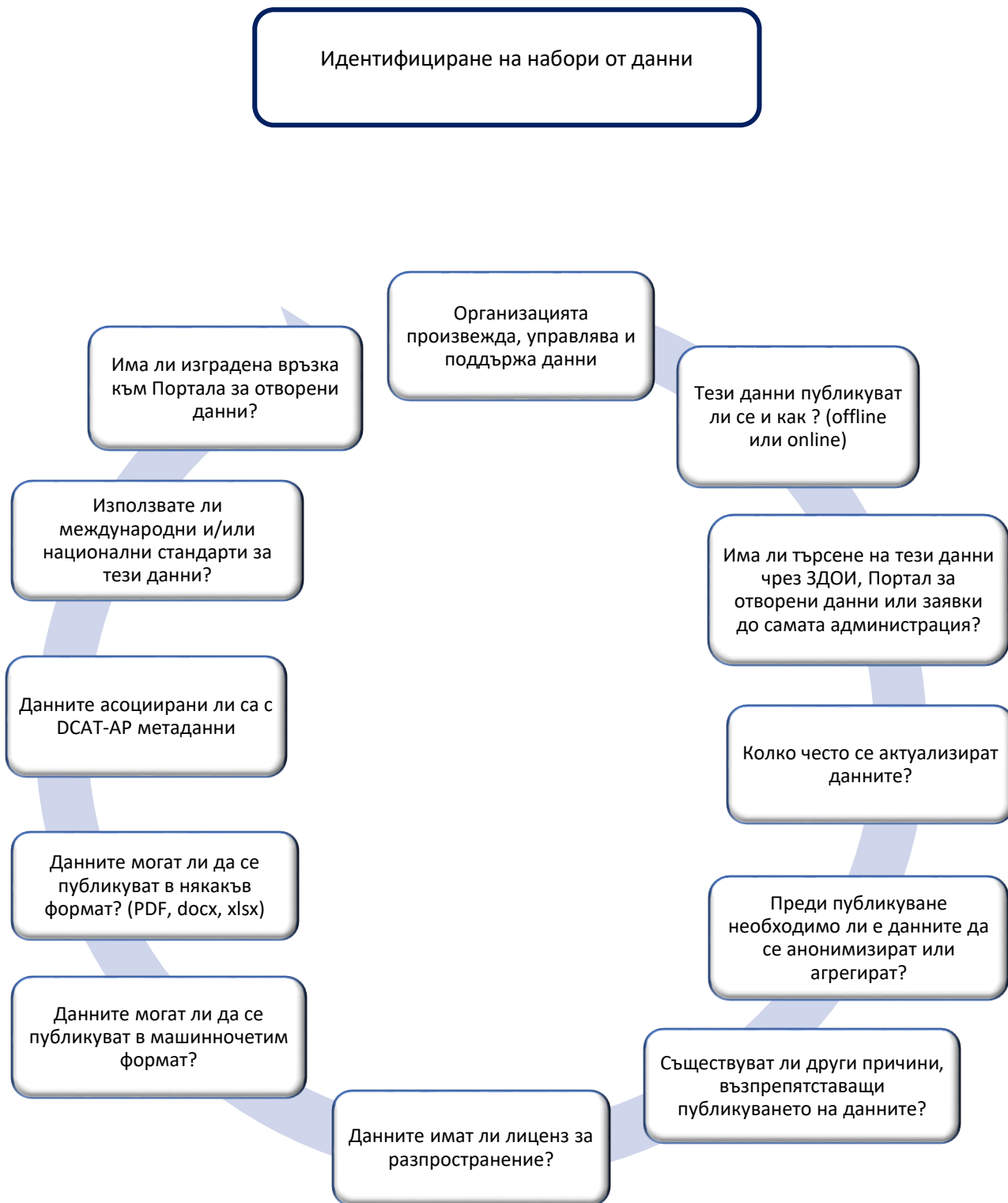
Техническата рамка се състои от следните основни елементи:

- Стандартни условия за повторно използване на информацията от обществения сектор и за нейното публикуване в отворен формат (стандартни лицензи);
- Формати за отворени данни;
- Схема на метаданни;
- Стандарти за отворени данни;
- Уникални адреси за идентификация на ресурсите (Uniform Resource Identifiers).

5.1. Процес за вземане на решение за публикуване на отворени данни в Портала за отворени данни.

Техническата рамка определя планиран и структуриран подход за публикуване на масиви от данни като отворени данни. Публичната администрация, при вземане на решение за публикуването на информация във вид на отворени данни, следва да вземе предвид *стойността, потенциала за повторно използване* на данните и *приноса им за постигане на по-добри резултати* за гражданите, бизнеса и другите държавни институции и *подобряване процеса на вземането на решения* в публичния и частния сектор.

Пример за процеса за вземане на решение за публикуване на данни в отворен формат, с включени основните въпроси, на които администрациите трябва да имат отговор, е показан на Фигура 1.



Фигура 1: Пример за процеса на вземане на решение за публикуване на данни в отворен формат

5.2. Оценка на отворените данни.

За да може да се направи оценка на публикуваните данни се използва пет степенна скала, препоръчана от Тим Бърнърс Лий. Тя е полезен инструмент за класифициране на ресурсите от гледна точка на тяхната отвореност. По-високата оценка по тази скала означава по-лесна последваща машинна обработка на данните и се използва за измерване степента на повторно използване и потенциалните ползи и въздействие. Скалата за оценка на отворените данни е описана в Таблица 1.

Ниво	Оценка	Формат	Допълнителна информация
1	★	Данните са налични онлайн, независимо от формата, под отворен лиценз.	Данните могат да са публикувани под формата на графични файлове, сканирани данни и текст или текстов файл.
2	★★	Данните са налични онлайн в машинночетим, структуриран вид;	Данните са публикувани под формата на форматирани таблици.
3	★★★	Данните са налични онлайн в машинночетим, структуриран вид, в отворен формат и не е необходимо да се използват платени лицензи;	Данните са публикувани в отворен, структуриран формат – CSV или JSON
4	★★★★	Данните са налични онлайн в машинночетим, структуриран вид, публикувани са в отворен формат и притежават уникален идентификатор (URI), за да са лесно достъпни и откриваеми от всяка точка в глобалната мрежа;	Данните са публикувани във формат, който разрешава да се използват етикети за семантичната структура на данните.
5	★★★★★	Данните са налични в машинночетим, структуриран вид, публикувани са в отворен формат, притежават уникален идентификатор (URI), за да се обозначат и са свързани с други данни в определен контекст.	Данните включват онлайн структурни връзки към други набори от данни.

Таблица 1. Скала за оценка на отворените данни.

5.3. Стандартни условия за повторно използване на информацията от обществения сектор и за нейното публикуване в отворен формат (стандартни лицензи).

Стандартните общи условия за повторно използване са въведени в българската нормативна уредба чрез Наредба за стандартните условия за повторно използване на информацията от обществения сектор и за нейното публикуване в отворен формат на Министерския съвет от юни 2016 година. В наредбата се въвежда и терминът „набор от

данни“ – еднороден тип информация – файлове в отворен формат с едно и също предназначение и структура. Те трябва да отговарят на изискванията на чл. 4 – 6 от Наредбата за стандартните условия за повторно използване на информацията от общественения сектор и за нейното публикуване в отворен формат. Към тях именно се отнасят стандартните общи условия.

Съгласно Наредбата за стандартните условия за повторно използване на информацията от общественения сектор и за нейното публикуване в отворен формат, организациите от общественения сектор предоставят информация с права за свободно повторно използване, преработка и разпространение при следните стандартни условия:

- Условия за предоставяне на информация без защитени авторски права (Приложение № 1 към чл. 15, ал. 1);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права (Приложение № 2 към чл. 16, ал. 1);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права. Споделяне на споделеното (Приложение № 3 към чл. 17, ал. 1, т. 1);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права. Без производни и сборни произведения (Приложение № 4 към чл. 17, ал. 1, т. 2);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права. Без употреба за търговски цели (Приложение № 5 към чл. 17, ал. 1, т. 3);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права. Без употреба за търговски цели. Споделяне на споделеното (Приложение № 6 към чл. 17, ал. 1, т. 4);
- Условия за предоставяне на произведение за повторно използване. Признаване на авторските права. Без употреба за търговски цели. Без производни и сборни произведения (Приложение № 7 към чл. 17, ал. 1, т. 5);
- Ограничения по отношение на отделни части на информацията, съдържащи производствена или търговска тайна (Приложение № 8 към чл. 20, ал. 1).

Следва да се има предвид, че в случаите, в които информацията не съдържа и не е съпроводена от съобщение или от условия за повторно използване, може да се използва без условия и ограничения.

За организациите от общественения сектор, които публикуват на Портала за отворени данни информация, е предвидена възможността да включват съобщение под формата на текстово съобщение, контекстуален прозорец или препратка към външен уебсайт, което да указва какъв е режимът на ползване на конкретния набор от данни. По този начин се посочва правният статут на информацията, като той се определя основно от наличието или не на авторски права.

5.4. Препоръчвани формати за отворени данни.

Докато публикуваните в произволен формат данни могат да се считат за отворени данни, то ако са свързани с отворен лиценз, използваният за тяхното представяне формат може да има определящо значение до колко могат да се използват. В рамките на политиката за отворени данни публичната администрация следва да публикува своите данни по възможно най-използваемия и подходящия начин. За да се определи до колко

даден формат е подходящ за публикуване се използва скалата за оценка на отворените данни, описана в Таблица 1.

За да могат данните да бъдат публикувани на Портала за отворени данни, публичните администрации трябва да изберат поне един от форматите, притежаващи поне 3 звезди по скалата за отворени данни. Въпреки това се насърчава публикуването на масиви от данни в няколко формата, като допълнение към препоръчаните 3 звезди (CSV).

Препоръки:

Всички набори от данни на <https://data.egov.bg/> трябва да са налични в поне един от следните формати:

General	Geospatial	Domain-Specific
CSV	GeoJSON	NetCDF
JSON	GML	Datex II
XML	KML	GTFS
ODF	WKT	JSON STAT
RDF	IFC	
ODS		

Промени във формата на публикуваните данни следва да се извършват след предварително уведомяване на потребителите, чрез посочване на датата, от която ще бъде въведен новият формат. Необходимо е да се даде достатъчно време (поне три месеца), за да се позволи на потребителите да вземат всички необходими мерки и да се гарантира, че няма да бъдат засегнати от промяната.

Масивите от данни могат да бъдат публикувани в няколко формата.

За масивите от данни, които все още не са налични в един от препоръчаните отворени формати, трябва да се разработи график със срокове за публикуването им в отворен формат.

Публикуването на данни в отворени формати следва да бъде вградено в процесите на публикуване на данни в публичния сектор и като част от управлението на информацията по-общо.

5.5. Използване на схема за метаданни в областта на отворените данни.

Порталът за отворени данни предоставя каталог на метаданни, чрез които се осъществява достъп до данни както от страна на национални потребители, така и от институциите и органите на ЕС. За да се улесни повторното използване на информацията, тези метаданни трябва да се основават на общи правила за кодиране и стандартизирани речници.

За да се подпомогне реализирането на ползите от отворените данни, публичната администрация следва да направи данните си по-лесни за търсени и използване. Необходимо е да се предоставят точни описания на масивите, за да бъде максимално улеснено идентифицирането, откриването на местоположението и онлайн извличането на информационните ресурси от потребителите.

Тези описания на данните се наричат „метаданни“.

Метаданните са обобщената информация, която описва наличността, естеството и съдържанието на данните. Те предоставят контекст, който помага на потребителите да разберат значението на данните, като например:

- наименование на набора от данни;
- за какво се отнася;
- къде може да се намери наборът от данни;
- кога е произведен и кога е последната му актуализация;
- какъв е източникът на данните;
- има ли някакви ограничения за използването на данните;
- методология на производство на данните.

Политиката за отворени данни изисква последователен подход към публикуването на отворени данни, за да се осигури оперативна съвместимост между масивите, публикувани от публичния сектор, както на национално, така и на международно равнище.

Към настоящия момент, съгласно Наредбата за стандартните условия за повторно използване на информация от обществения сектор и за нейното публикуване в отворен формат, организациите от обществения сектор публикуват на портала за отворени данни информационните масиви и ресурси, заедно със съответните метаданни, предвидени при разработката на Портала. Използваната схема за метаданните е съгласно проекта „Подобряване на процесите, свързани с предоставянето, достъпа и повторното използване на информацията от обществения сектор“, в рамките на който е реализиран Портала за отворени данни.

За да се добави нов набор от данни е необходимо да се попълни формуляра, който се визуализира на страницата за добавяне на набор от данни. Само полетата, маркирани със символа * (звездичка) са задължителни. Останалите може да се попълнят, както при създаването на новия набор от данни, така и в последствие, чрез опцията за редактиране. Формулярът включва следните полета:

- **„Уникален идентификатор“** – в това поле може да се попълни идентификатор на набора данни, ако например е необходимо той да отговаря на конкретна номенклатура. Ако такъв не бъде попълнен, системата автоматично присвоява такъв, който може да се редактира в последствие;

- **„Основна тема“** – това поле е маркирано като задължително. От падащото меню може да се избере основна тема. Ако наборът от данни не принадлежи към никоя от посочените, може да се избере опцията „Некатегоризирани“;

- **„Наименование“** – кратко представяне на същността на публикуваните данни или нормативно установеното наименование;
- **„Описание“** – по-детайлно представяне на същността на публикуваните данни;
- **„Споразумение на ниво обслужване“** – информация за: период на обновяване; времето, в което структурите на данните остават непроменени;
- **„Етикети“** – подтеми, за които се отнасят данните в набора (напр. за тема „здравеопазване“, подтеми: аптеки, болници, лекарства, цени, НЗОК и т.н.). Всеки един етикет, който се присвоява на набора от данни го прави по-лесен за намиране при търсене с ключова дума.
- **„Условия за ползване“** – съгласно Наредбата за стандартните условия за повторно използване на информацията от обществения сектор и за нейното публикуване в отворен формат условията за ползване при публикуване на данни включват изброените в падащото меню опции. Ако трябва да се публикува набор от данни при различни от посочените условия, може да се направи заявка до Администратор на портала за нови условия, като се избере бутон „Нови условия“. В прозореца, който се появява, може да се попълни в полето „Описание“ условията за ползване, които се иска да бъдат добавени. Останалата информация се попълва автоматично;
- **„Организация“** – име на организацията, от чието име се публикуват данните, ако има такава;
- **„Група“** – име на групата, към която се присвоява съответния набор от данни;
- **„Видимост“** – възможност за определяне на „публичност“ на публикуваните данни на портала. В случай, че публикуваната информация е конфиденциална, системата дава възможност за отбелязване на този факт и респективно ограничаване на достъпа до съответната информация, чрез опцията „Частен“ в падащото меню;
- **„Източник“** – източник на данните в случай, че данните вече са публикувани онлайн;
- **„Автор“** – юридическо или физическо лице, което е създадо или събрало данните;
- **„E-mail на автора“** - адрес на електронната поща на автора;
- **„Контакт“** – Лицето, което отговаря за поддържането на данните в актуално състояние;
- **„E-mail за контакти“** – адрес на електронната поща на лицето, което отговаря за поддържането на данните в актуално състояние;
- **„Допълнително поле“** – поле/полета, в които може да се попълва допълнителна информация, която не присъства в нито едно от останалите полета (напр. нормативен акт или др.).

Забележки:

- при избор на Група или Организация от падащите менюта се визуализират само групите и организациите, в които членува лицето, което създадо набора от данни;
- всеки нов набор от данни остава в режим „чернова“ и е видим само за лицето, което го е създадо. За да бъде публикуван, към него трябва да се добави поне един ресурс. След качване на ресурс или ресурси в набора, публикуването се осъществява чрез избор на бутон „Публикувай“, който е достъпен в самия набор от данни.

В набор от данни може да се качи ресурс, като се избере бутон „Добави ресурс“, след което се зарежда страница за добавяне на ресурс към същия набор от данни. Само полетата маркирани със символа * (звездичка) са задължителни. Останалите може

да се попълнят, както при създаването на новия ресурс, така и в последствие, чрез опцията за редактиране. В страницата се попълват метаданни в следните полета:

- **„Наименование“** – кратко представяне на ресурса;
- **„Описание“** – по-детайлно представяне на същността на данните;
- **„Тип“** – тип на ресурса. Възможни стойности: *API, Файл, Хиперлинк, Автоматично качване*. След избор на тип на ресурса, се визуализират различни полета за въвеждане на данни.

○ **За тип „API“:**

- **„Връзка“** - препратка (URL адрес) към услугите на интерфейса;
- **„Тип на заявката“** - тип на заявката, която трябва да се изпрати - POST, GET;
- **„HTTP хедъри“** - ако е необходимо изпращане на хедъри, те трябва да се въведат в това поле;
- **„Заявка“** - заявка към услугата;
- **„Тип на честота на обновяване“** - тип на честотата на обновяване - час, ден, седмица, месец;
- **„Честота на обновяване“** - стойност на честотата на обновяване.

○ **За тип „Файл“** – **„Файл“** - поле за прикачване на файл. Прикаченият файл може да бъде в отворен формат като: CSV, JSON, KML, RDF, XML, HTML, TXT, XSD. Възможно е прикачване на файлове от затворен тип: XSL/XSLX, DOC/DOCX, Изображения (сканирани документи) - jpg; svg; png, PDF (сканиран), PDF (експортиран). При прикачване на файлове от затворен тип, автоматично се стартира конвертора на данни, който при възможност извлича текстовата информация от прикачения файл.

○ **За тип „Хиперлинк“** – **„Връзка“** - препратка (URL адрес) към страницата, на която за публикувани данните.

○ **За тип „Автоматично качване“** не е необходимо въвеждането на допълнителна информация. Този тип се задава за ресурси, чиито данни се качват и обновяват чрез инструмента за автоматично качване на данни на Портала.

Общи полета за всички типове ресурси:

- **„Описание на схема“** – съдържа информация за данните в ресурса;
- **„Връзка към схема“** – препратка към страница, на която е публикувана схемата на данните;
- **„Допълнително поле“** – поле/полета, в които може да се попълва допълнителна информация, която не присъства в нито едно от останалите полета.

5.6. Използване на стандарти за отворени данни.

Стандартите за данни, също наричани модели на данни или информационни речници, осигуряват общо разбиране на съдържанието и описанието на данните. Чрез стандартите се улеснява обменът и се гарантира публикуването на отворените данни по постоянен/устойчив и последователен начин.

Стандартите за данни помагат да се даде общо значение на данните. Това е особено важно, когато те се използват от трета страна, когато са интегрирани от различни източници, или когато се споделят. Стандартите за данни не само дефинират значението на определени понятия, но също така и начина, по който понятията се отнасят помежду си, което улеснява оперативната съвместимост на данните.

Когато се публикуват отворени данни, ако е възможно, трябва да се използват международните стандарти, дефинирани от признати и/или сертифицирани организации по стандартизация, като например ISO, W3C, IETF, OGC и OASIS. Ако международните

стандарти са недостъпни или неподходящи, се препоръчва да се използват национални такива. За конкретни теми като геопространствени, статистически или здравни, се препоръчва използването на национални стандарти, определени от съответно отговорната организация.

5.7. Уникални адреси за идентификация на ресурсите (Uniform Resource Identifiers).

Техническата рамка определя стъпките, които позволяват на администрациите да постигнат минимум трето ниво на отворени данни. В по-дългосрочен план обаче, намерението е да се постигне по-високо ниво на отворените данни (с оценка 4 и 5 звезди).

Използването на Uniform Resource Identifiers (URIs) е важен елемент от дългосрочния подход към отваряне на данните.

Всеки набор от данни трябва да притежава уникален адрес за идентификация, който представлява последователност от символи, идентифицираща абстрактен или физически ресурс. На този етап в Портала за отворени данни има функционалност, чрез която системата автоматично присвоява уникален идентификатор, ако такъв не бъде посочен от организациите при публикуване на набор от данни или ресурс.

6. Актуализиране, поддържане и достъпност на данните и ограничения за публикуване.

6.1. Задължително поддържащо актуализиране на публикуваните данни.

Отворените данни трябва да се предоставят/публикуват възможно най-бързо след като бъдат събрани. Еднократно публикуван набор от данни като отворен става непълен веднага щом се съберат, но не се публикуват, допълнителни данни към набора.

Проблемът с бързото публикуване на данни може да бъде разрешен като се комбинират някои мерки за дигитализиране на събирането и публикуването на данни (напр. електронно подаване на информацията, централизирани хранилища за данни и API).

6.2. Създаване на постоянен, устойчив достъп до данните.

Веднъж публикувани, цифровизираните данни трябва да останат постоянно достъпни и „откриваеми“ на постоянно място в мрежата (постоянни URI) или в постоянен архив. Това изискване важи както за порталите и уебсайтовете, които са средство за достъп до тези данни, така и за самите набори от данни.

За да бъдат разбираеми от потребителите, всички промени и/или актуализации в данните трябва да бъдат подробно документирани, като включват подходящо проследяване на версиите и архивиране във времето.

6.3. Ограничения при разпространението на публичните данни като отворени.

Някои данни, които административните органи събират и съхраняват, не са подходящи за разпространение. Недопустимо е публикуването на данни, което би довело до нарушаване на правото на неприкосновеност на личния живот или които съдържат класифицирана информация. Националният статистически институт също може да откаже публикуването на определени данни с цел опазване на статистическата тайна.

Изключително важно е да се гарантира неприкосновеността на личния живот, като в същото време е необходимо да се избере подходяща техника (напр. анонимизиране), за да се запази информационната стойност на даден набор от данни и неговият потенциал за повторна употреба.

Липса на законова обосновка

Основната причина, която възпрепятства използването на публични данни, е липсата на подходящи правни регламенти по отношение на отвореността на данните, което води до много усложнения за потенциалния потребител на такива данни.

Най-често срещани пречки за предоставянето на данни:

- чрез затворени или ограничени формати, които възпрепятстват експортирането към отворен файлов формат;
- за които няма информация за датата на създаване или актуализиране;
- които са обект на защита на авторското право;
- които са обвързани с лиценз, чиито условия забраняват неограничен обмен.

Лицензионни ограничения

Лицензионните ограничения са правните ограничения пред използването на публични данни. Всички публикувани данни трябва да бъдат обвързани с лиценз, който регламентира правата за използване на данните.

Задължителната информация, която лицензът трябва да предоставя е: наименование на автора, вид на лиценза, заглавие на произведението и връзка към източника на данните.

Защита на авторски права

Регулациите за защита на авторските права могат да бъдат пречка за свободното използване на данни. Тази защита трябва да се разглежда както в контекста на споделянето на данните, така и в регистъра, в който са налице данните.

Данните, които имат атрибути на произведение, чиито авторски права ще бъдат предоставени на трети страни, не могат да бъдат използвани повторно.

Данните и метаданните, които са предмет на авторски права, трябва да бъдат ясно обозначени/маркирани от притежателя им.

Тази практика ще позволи на потенциалните потребители бързо да получат информация за авторските права на данните и метаданните. Предоставянето на информация за лицето, което има право на икономически авторски права, ще позволи на заинтересованите потребители да предприемат необходимите действия, за да получат разрешението на автора за повторно използване на данните и метаданните.

Заплащане при разпространение на данните

Друга бариера пред разпространението на информация е налагането на такси за споделяне на данни.

Препоръчително е безплатното предоставяне на данни да се приема като принцип. Таксите могат да се събират само когато притежателят на данните е направил допълнителни разходи, свързани с подготовката им за предоставяне по начин, заявен от потребителя.

Наличие на данните в затворен или ограничителен/рестриктивен формат

Публикуването на данни във формати, които не са машинночетими или изискват наличие на платен софтуер, е основна пречка за използването на данните. Това означава, че спецификацията на форматите на данните е собственост на компаниите, които създават тези формати, а възможността за търсене, анализ и повторно използване на такива данни изисква значителни разходи за персонал и финансови средства (например, ако е необходимо да се закупи специализиран софтуер).

Препоръчва се споделянето на публичните данни да бъде в поне един отворен формат, чиито свойства са публични и добре документирани. Тази препоръка може да се реализира и чрез експорт на данните в отворен формат.

Защита на личните данни

Ограничение за разпространението на данните като отворени възниква от действия на техните притежатели, предприети за защита на неприкосновеността на личните данни на физическите лица. В някои случаи обаче разпоредбите на закона се използват като средство за отказ от предоставянето на данни или за предоставянето им, но по начин, който е непълен от гледна точка на потребителя, т.е. в нарушение на принципа за пълнота на отворените данни.

Съгласно разпоредбите на Закона за защита на личните данни, предоставянето на данни, които включват лични данни, за последващо повторно използване, подлежи на задължително прилагане на принципите за защита на личните данни. Предоставянето на такива данни изисква предварителна обработка (анонимизация) от притежателя на данните, т.е. трябва да бъдат предприети действия за премахване на всички данни, които дават възможност за идентифициране на конкретно физическо лице в набора от данни.

Степента, до която данните трябва да бъдат анонимизирани, трябва да бъде внимателно анализирана, така че да не се ограничава достъпът и последващото повторно използване на данните.

В изключителни случаи предоставянето на данни може да бъде отказано, когато анонимизацията не гарантира подходяща защита на личния живот на лицето, за което се отнасят тези данни.

Администраторите на лични данни задължително спазват ограниченията, въведени с Регламент (ЕС) 2016/679 на Европейския парламент и на Съвета от 27 април 2016 година относно защитата на физическите лица във връзка с обработването на лични данни и относно свободното движение на такива данни и за отмяна на Директива 95/46/ЕО Общ регламент относно защитата на данните.

7. Практически насоки за вътрешни правила в организациите от общественения сектор с цел организиране на дейностите в областта на отворените данни.

7.1. Приложение на принципите за отворените данни.

За реално разпространение на отворени данни, а следователно и за ползите и възможността да се използват отворени данни, можем да говорим само, ако те се предоставят в съответствие с основните принципи за отворени данни.

Във всяка администрация е необходимо да се разработят правила, свързани с производството и разпространението на отворени данни.

Пример за вътрешни правила:

- Публичните данни, събрани в регистър, са общодостъпни, безплатни и с право за повторна употреба без ограничения.
- Осигурен е достъп до публичните данни, които са:
 - актуални;
 - пълни;
 - във вида на предоставяне от източника;
 - в машинночетим формат.
- Прилагат се подходящи технически и организационни мерки за осигуряване на достъп до данните в съответствие с горепосочените правила за достъп и разпространение.

7.2. Актуализиране на данните.

Препоръчва се да се въведат вътрешни правила относно интервалите за актуализиране на данните, така че потребителите да са запознати с актуалността на данните и съответно да бъде показано кога се очаква следващото обновяване.

Пример за вътрешни правила:

- Публичните данни се актуализират напр. веднъж седмично.
- Времето за актуализиране трябва винаги да съответства на вида на данните, които се предоставят.

7.3. Споделяне на данни чрез уеб услуги.

Публичните данни могат да се предоставят и чрез уеб услуги. Препоръчително е да се въведе автоматизирано решение, което да осигури достъп до данните на организацията (търсене, разглеждане, извличане и обработване) чрез т.нар. уеб услуги. Тези услуги предлагат предимството на предоставянето на данни, независимо от софтуерната и хардуерна платформа на потребителя, което улеснява оперативната съвместимост и използването на данни.

Пример за вътрешни правила:

- Методи и правила за използването и предоставянето на данни чрез уеб услуги.

7.4. Определяне на длъжностно лице или лица, отговорни за политиката за отворени данни в рамките на организацията от общественния сектор.

За да се гарантира правилното прилагане на принципите за отворени данни, всеки собственик/производител на данни е необходимо да определи и/или назначи длъжностно лице, отговорно за политиката за отворени данни. При необходимост да се създадат или променят съществуващи вътрешноведомствени разпоредби, с които да се регламентират техните права и задължения.

За да се даде възможност на тези служители да работят успешно и ефективно, се препоръчва да се въведат подходящи вътрешни правила и процедури.

Пример за вътрешни правила:

- Определя се длъжностно лице, което отговаря за отворените данни.

- Отговорностите на служителя включват наблюдение на спазването на правилата за публикуване, разпространение и предоставяне на данните като отворени, като:

- оценява качеството и полезността на публичните данни;
- отговаря за посочване на данни от особено значение за развитието на иновациите в страната и за развитие на информационното общество, които да се предоставят за повторна употреба;
- служи като контактна точка за потребителите на данни по въпроси, свързани с достъпа до обществена информация и повторното използване на информация от обществеността;
- изготвя доклади за прилагане на препоръчаните мерки в областта на отворените данни в публичния сектор.

8. Начини за публикуване на информация на Портала за отворени данни.

В резултат на реализацията на проекта „Подобряване на процесите, свързани с предоставянето, достъпа и повторното използване на информацията от обществеността“, насочен към процесите, свързани с поддържането и публикуването на публичната информация в отворен формат за повторна употреба, както и към създаване на нов механизъм за заявяване и получаване на достъп до обществена информация по електронен път и публикуването ѝ в интернет, са изградени Портал за отворени данни и Платформа за достъп до обществена информация. Те са предпоставка за улеснено повторно използване на обществената информация, чрез публикуването ѝ в отворен формат и за подобряване на механизмите за достъп до информацията.

Новият Портал за отворени данни дава възможност за публикуване на данни по следните начини:

- чрез ръчно въвеждане на набори от данни през интерфейса на портала;
- чрез приложно-програмен интерфейс (API), реализиран като уеб услуга, позволяващ качване и обновяване на данни. Програмният интерфейс предоставя всички данни от портала в отворен, машинно четим формат, подходящ за директна повторна употреба;
- чрез препратка (link) към външна система на организация от обществеността, където са публикувани данните.

Публикуването на информация на Портала за отворени данни се извършва съгласно разработеното в рамките на проекта Ръководство за публикуване на данни на портала.