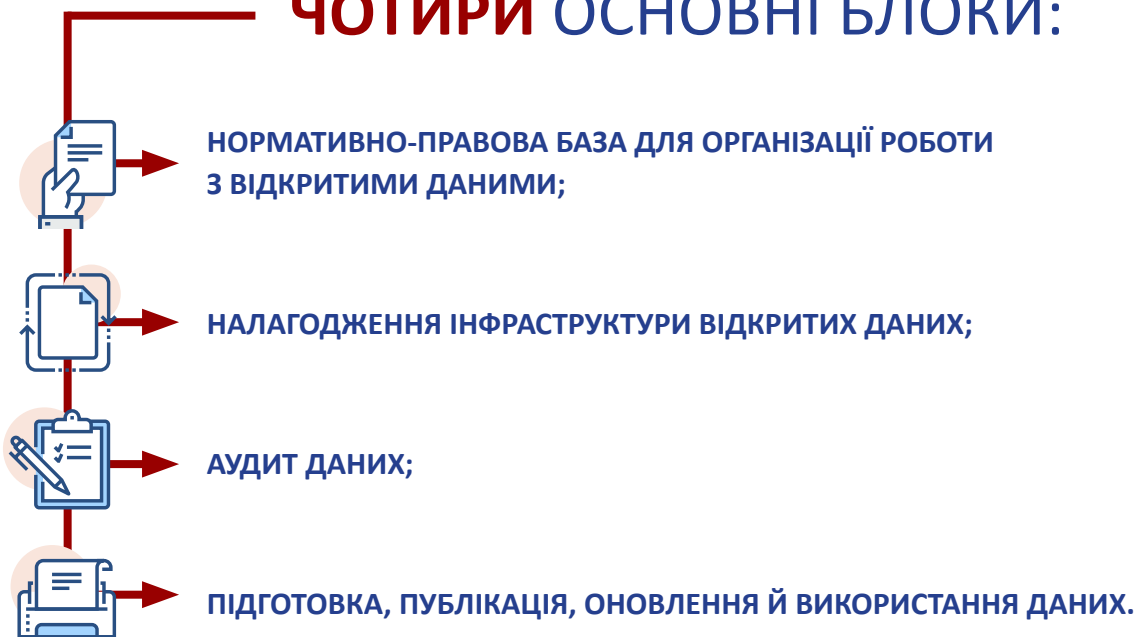


Методичні рекомендації підготовлені
Державним агентством з питань електронного урядування України
з залученням експертів програми
“Прозорість та підзвітність у державному управлінні та послугах”.

МЕТОДИЧНІ РЕКОМЕНДАЦІЇ ОХОПЛЮЮТЬ

ЧОТИРИ ОСНОВНІ БЛОКИ:



ВСТУП ДО ВІДКРИТИХ ДАНИХ

ЩО ТАКЕ ВІДКРИТІ ДАНІ

Публічна інформація у формі відкритих даних (відкриті дані) – це дані, які будь-хто може вільно використовувати й поширювати.

Відкриті дані **мають бути опубліковані** в мережі Інтернет у машиночитаному форматі – форматі, придатному для автоматизованого оброблення інформаційними системами.

Влада, бізнес, громадські організації, наукові установи і приватні особи можуть **вільно публікувати, копіювати, поширювати, використовувати відкриті дані** для одержання економічної, соціальної, культурної чи екологічної вигоди.

Відкриті дані можна **вільно використовувати**, зокрема з комерційною метою, у поєднанні з іншою інформацією або включаючи їх до складу власного продукту з обов'язковим зазначенням джерела їх отримання.

ОЗНАКИ І ПРИКЛАДИ ВІДКРИТИХ ДАНИХ

Найкраща практика оприлюднення відкритих даних:

- Первинні, неагреговані, гранулярні та максимально деталізовані дані.
- Машиночитаний формат, який дає змогу інформаційним системам ідентифікувати, розпізнавати, перетворювати й отримувати конкретні дані без участі людини. Поширені машиночитані формати – CSV, XML, JSON.
- Дані у вільному доступі для використання й поширення, зокрема з комерційною метою.

ПРИКЛАДИ:



Дані [Єдиного реєстру юридичних осіб та фізичних осіб-підприємців](#) доступні на Єдиному державному веб-порталі відкритих даних у форматі XML. Розпорядник – Міністерство юстиції України. Періодичність оновлення – не пізніше ніж за 5 робочих днів після внесення змін.



[Єдиний реєстр судових рішень за 2018 рік](#) у форматі CSV. Розпорядник – Державна судова адміністрація України.



[Перелік транспортних засобів](#), що перебувають у розшуку у зв'язку з їх незаконним заволодінням, у форматі JSON. Розпорядник – Міністерство внутрішніх справ України.



Відкриті дані, які доступні через інтерфейс прикладного програмування (API), зокрема машиночитані дані [Єдиного порталу публічних фінансів](#). Розпорядник – Міністерство фінансів України.

ЩО НЕ Є ВІДКРИТИМИ ДАНИМИ

- Оприлюднення структурованої публічної інформації (звітів, планів та інших документів) у немашиночитаних форматах на офіційних сайтах. Наприклад, оприлюднення таблиць у відсканованому вигляді не є відкритими даними.
- Електронні сервіси без доступу до вихідних первинних даних, зокрема інтерактивні мапи, електронні черги, петиції, документообіг тощо.

ПРИКЛАДИ:



Сервіс "[Перевірити вашого бізнес-партнера](#)" Державної фіскальної служби не є відкритими даними, хоча він побудований на основі даних.



[Інтерактивна мапа доріг](#) Державного агентства автомобільних доріг України, яка надає інформацію про відрізки доріг, де відбувається ремонт чи є обмеження руху. Однак доступу до первинних даних, на яких ця мапа побудована, немає.

ПОЧАТОК РОЗВИТКУ ВІДКРИТИХ ДАНИХ В УКРАЇНІ

В Україні відкриті дані вперше офіційно згадуються в Національній антикорупційній [стратегії](#) на 2014-2017 роки та в [Плані](#) дій із впровадження Ініціативи “Партнерство “Відкритий Уряд” у 2014-2015 роках. У цих документах йшлося про необхідність створення нормативно-правових та організаційних основ для запровадження доступу до інформації у формі відкритих даних та її повторного використання. Згодом Верховна Рада України спільно з громадськістю розробила та ухвалила зміни до Закону України “Про доступ до публічної інформації”, які визначили поняття публічної інформації у формі відкритих даних та особливості її використання.

ВПРОВАДЖЕННЯ ВІДКРИТИХ ДАНИХ В УКРАЇНІ

За реалізацію державної політики у сфері відкритих даних [відповідає](#) Державне агентство з питань електронного урядування України, яке координує Кабінет Міністрів України. Цей центральний орган виконавчої влади забезпечує функціонування Єдиного державного веб-порталу відкритих даних (<http://data.gov.ua/>), який надає доступ до публічної інформації у формі відкритих даних.

Державне агентство з питань електронного урядування розробляє та впроваджує стратегічні документи, рекомендації щодо відкритих даних, надає підтримку й консультації розпорядникам інформації, а також працює над удосконаленням роботи Єдиного державного порталу відкритих даних України, якістю наборів даних та їх використанням.



НОРМАТИВНО-ПРАВОВА БАЗА ДЛЯ ОРГАНІЗАЦІЇ РОБОТИ З ВІДКРИТИМИ ДАНИМИ

НАЦІОНАЛЬНЕ ЗАКОНОДАВСТВО

Закон України “Про доступ до публічної інформації”

В Україні відкриті дані регламентує [Закон України “Про доступ до публічної інформації”](#). Цей Закон визначає поняття відкритих даних (стаття 10.1) та умови їх використання (стаття 10.2).

Стаття 10.1. **Публічна інформація у формі відкритих даних** – це публічна інформація у форматі, що дозволяє її автоматизоване оброблення електронними засобами, вільний та безоплатний доступ до неї, а також її подальше використання.

Розпорядники інформації зобов’язані надавати публічну інформацію у формі відкритих даних на запит, оприлюднювати і регулярно оновлювати її на єдиному державному веб-порталі відкритих даних та на своїх веб-сайтах.

Стаття 10.2. Публічна інформація у формі відкритих даних **є дозволеною для її подальшого вільного використання та поширення**.

Будь-яка особа може вільно копіювати, публікувати, поширювати, використовувати, у тому числі в комерційних цілях, у поєднанні з іншою інформацією або шляхом включення до складу власного продукту, публічну інформацію у формі відкритих даних з обов’язковим посиланням на джерело отримання такої інформації.



ПОЛОЖЕННЯ ПРО НАБОРИ ДАНИХ, ЯКІ ПІДЛЯГАЮТЬ ОПРИЛЮДНЕННЮ У ФОРМАТІ ВІДКРИТИХ ДАНИХ

Для виконання Закону України “Про доступ до публічної інформації” Кабінет Міністрів України розробив та ухвалив [Постанову №835 “Про затвердження Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних”](#), яка визначила формати, структуру, вимоги, перелік наборів відкритих даних, що обов’язкові до оприлюднення розпорядниками публічної інформації.

20 грудня 2017 року Кабінет Міністрів України ухвалив зміни до Постанови №835, розширивши перелік наборів для оприлюднення ([Постанова Кабінету Міністрів України від 20.12.2017 №1100](#) “Про внесення змін до Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних”).

ПОРЯДОК ВЕДЕННЯ ЄДИНОГО ДЕРЖАВНОГО ПОРТАЛУ ВІДКРИТИХ ДАНИХ УКРАЇНИ

Робота Єдиного державного порталу відкритих даних України та взаємодія розпорядників інформації з ним затверджені [Постановою Кабінету Міністрів України №867 “Деякі питання оприлюднення публічної інформації у формі відкритих даних. Порядок ведення Єдиного державного веб-порталу відкритих даних”](#).

МІЖНАРОДНА ХАРТІЯ ВІДКРИТИХ ДАНИХ ТА УКРАЇНА

[Міжнародна хартія відкритих даних](#) – це ініціатива щодо співпраці урядів, приватного сектору та неурядових організацій для розвитку відкритих даних у світі. Вона була заснована у 2015 році. Наразі до цієї ініціативи долучилися більш ніж 70 урядів та організацій.

Україна долучилася до Міжнародної хартії відкритих даних 22 вересня 2016 року. Кабінет Міністрів України ухвалив [Розпорядження №686-р “Деякі питання приєднання до Міжнародної](#)



[Хартії відкритих даних](#)”, яке визначило принципи роботи держави у сфері розвитку відкритих даних.

З текстом Хартії відкритих даних українською мовою можна ознайомитися [за посиланням](#).

Хартія містить основні принципи:

- Відкритість за замовчуванням;
- Чіткість і оперативність;
- Доступність і використання;
- Порівнюваність та інтеперабельність;
- Розвиток електронного урядування та залучення громадян;
- Інклюзивний розвиток та інновації.

Державне агентство з питань електронного урядування співпрацює з Міжнародною хартією відкритих даних задля розвитку відкритих даних в Україні відповідно до міжнародних стандартів. Зокрема, робота уряду у сфері відкритих даних базується на принципах, що закріплені в Міжнародній хартії відкритих даних.

ДОРОЖНЯ КАРТА РОЗВИТКУ ВІДКРИТИХ ДАНИХ В УКРАЇНІ

План роботи уряду щодо розвитку відкритих даних визначає Дорожня карта, яку розробляють і затверджують раз на рік із залученням експертів та активістів. 2017 року Дорожня карта розвитку відкритих даних в Україні була розроблена на основі принципів Міжнародної хартії відкритих даних. Вона визначила 59 завдань урядові щодо впровадження політики відкритих даних. На основі них були розроблені Плани дій для реалізації принципів Міжнародної хартії відкритих даних у п’яти містах: [Львів](#), [Вінниця](#), [Чернівці](#), Дрогобич, Дніпро.



ПЛАНИ ПРІОРИТЕТНИХ ДІЙ УРЯДУ

Розвиток відкритих даних – одне з першочергових завдань Кабінету Міністрів України, що визначено в середньостроковому [плані пріоритетних дій Уряду](#) до 2020 року та плані пріоритетних дій Уряду на 2017 рік. Зокрема, завданням до 2020 року є оприлюднення не менше ніж 50 000 наборів даних у формі відкритих даних і стимулювання розвитку суспільних та комерційних проектів на їх базі, зокрема не менше ніж 5 000 наборів і 10 проектів у 2017 році.

КОНЦЕПЦІЯ РОЗВИТКУ ЕЛЕКТРОННОЇ ДЕМОКРАТІЇ В УКРАЇНІ

[Концепція розвитку електронної демократії в Україні та План заходів з реалізації першого етапу її виконання](#), які затверджені 20 вересня 2017 року Розпорядженням Кабінету Міністрів України, визначають заходи щодо забезпечення розвитку відкритих даних до 2020 року.

У концепції зазначено, що оприлюднення публічної інформації у формі відкритих даних для вільного використання:

- дасть змогу створити додаткові механізми реалізації права на доступ до інформації, що перебуває у володінні органів влади;
- підвищить прозорість діяльності державних органів;
- забезпечить вільний обіг інформації та можливість її подальшого використання для реалізації особистих прав та свобод людини і громадянина, розвитку інновацій та стимулювання ведення господарської діяльності, а також запобігання та викриття корупції.

ОСНОВНІ ЗАХОДИ ДЛЯ ЗАБЕЗПЕЧЕННЯ РОЗВИТКУ ЕЛЕКТРОННОГО ВРЯДУВАННЯ В УКРАЇНІ

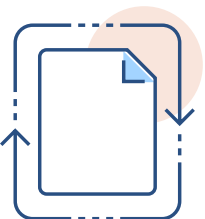
- Формування й розвиток інфраструктури відкритих даних на базі Єдиного державного порталу відкритих даних, інтегрованих із ним інших веб-порталів відкритих даних, публічних інтерфейсів прикладного програмування (API), відкритих стандартів та форматів.



- Оприлюднення й регулярне оновлення наборів даних у формі відкритих даних відповідно до суспільного інтересу, найкращих світових практик, встановлених вимог щодо якості даних, відкритості та прозорості діяльності.
- Стимулювання розвитку на базі відкритих даних загальнодоступних проектів і сервісів (соціальних, громадських, медійних і комерційних), зокрема у співпраці з органами влади, для підвищення відкритості та ефективності їхньої діяльності, надання якісних послуг та розвитку інноваційного бізнесу.

КОНЦЕПЦІЯ РОЗВИТКУ ЦИФРОВОЇ ЕКОНОМІКИ ТА СУСПІЛЬСТВА УКРАЇНИ НА 2018-2020 РОКИ

У проекті розпорядження Кабінету Міністрів України «Про схвалення [Концепції розвитку цифрової економіки та суспільства України на 2018-2020 роки](#) та затвердження плану заходів щодо її реалізації» відкриті дані визначені як один з напрямів розвитку цифрової економіки. У документі йдеться про ринковий потенціал відкритих даних для підвищення відкритості, прозорості й ефективності роботи державних установ, а також про шлях до розвитку нової для України “цифрової” індустрії – індустрії відкритих даних.



НАЛАГОДЖЕННЯ ІНФРАСТРУКТУРИ ВІДКРИТИХ ДАНИХ

СКЛАДОВІ ЧАСТИНИ ВПРОВАДЖЕННЯ ВІДКРИТИХ ДАНИХ

Для сталого, налагодженого та системного впровадження відкритих даних на державному та місцевому рівнях необхідні три взаємопов'язаних складники:

- **Нормативний** – впровадження нормативно-правової бази, процедур, які регламентують відкриті дані, перелік наборів даних для оприлюднення, роботу розпорядників, відповідальність, умови використання даних.
- **Організаційний** – сукупність кроків, які необхідно зробити органу державної влади для виконання завдань, поставлених у нормативно-правових актах.
- **Технічний** – передбачає те, що дані повинні бути оприлюднені в чітко визначених форматах, відповідно до стандартів, у певному місці (портал, репозитарій, сайт) та з відповідною частотою оновлення.

Згідно з найкращими міжнародними практиками, державні органи мають популяризувати, сприяти та заохочувати використання даних, зокрема для створення сервісів і продуктів. Успішні історії використання відкритих даних допомагають просувати культуру відкритих даних.



СТВОРЕННЯ ВНУТРІШНЬОЇ НОРМАТИВНОЇ БАЗИ ДЛЯ ОРГАНІВ ЦЕНТРАЛЬНОЇ ВЛАДИ

Окрім національного законодавства, яке визначає загальні принципи й підходи щодо роботи з відкритими даними, для врегулювання процесів на рівні органів виконавчої влади, місцевого самоврядування, державних і комунальних підприємств варто розробити додаткову нормативно-правову базу. Це необхідно для того, щоб забезпечити організаційну та інституційну підтримку впровадженню та розвитку відкритих даних, а також визначити відповідальних, їхнє коло повноважень, сферу роботи тощо.

Оприлюднення відкритих даних у **центральних органах виконавчої влади** регулюється внутрішніми наказами, які мають охоплювати такі питання:

- Визначення структурних підрозділів (секторів, відділів, департаментів) та/або осіб, відповідальних за роботу з відкритими даними;
- Затвердження порядку оприлюднення та оновлення наборів даних на Єдиному державному веб-порталі відкритих даних та на власних веб-сайтах.

ПРИКЛАДИ:

- Наказ [Міністерство Юстиції](#);
- Наказ [НАДС](#).

Як центральним органам виконавчої влади, так і органам місцевого самоврядування, доцільно ухвалити такі документи:

- Посадові обов'язки відповідальних осіб щодо відкритих даних;
- План дій із реалізації політики відкритих даних;
- Положення/порядок щодо відкритих даних.

Зразки вказаних документів є в наступному розділі.



СТВОРЕННЯ НОРМАТИВНОЇ БАЗИ ДЛЯ ОРГАНІВ МІСЦЕВОГО САМОВРЯДУВАННЯ

Оприлюднення відкритих даних **для органів місцевого самоврядування** може охоплювати такі завдання:

- Визначення відповідальних за відкриті дані структурних підрозділів (секторів, відділів, департаментів) та/або осіб;
- Затвердження порядку оприлюднення наборів даних на Єдиному державному порталі відкритих даних України, а за наявності також на власному порталі або сайті;
- Проведення та впровадження результатів аудиту даних;
- Створення за потреби робочої групи з впровадження та розвитку галузевих відкритих даних із залученням зацікавлених сторін.

[Зразки нормативних документів для міських рад](#)

[Зразок розпорядження/рішення виконкому](#)

Цей документ визначає основні завдання для розвитку політики відкритих даних у місті.

[Зразок Плану дій із реалізації політики відкритих даних](#)

Містить перелік заходів і відповідальних осіб/підрозділів за реалізацію політики відкритих даних на визначений період.



Зразок посадових інструкцій

Перелік прав та обов'язків відповідальної особи/органу за відкриті дані.

Зразок меморандуму про співпрацю щодо відкритих даних

Документ, що визначає мету й умови співпраці органів виконавчої влади, місцевого самоврядування та інших розпорядників даних.

Зразок Положення про відкриті дані

Зібраний документ щодо реалізації політики відкритих даних, що визначає основні поняття, принципи роботи, формати та стандарти.

ОРГАНІЗАЦІЯ ПРОЦЕСУ ВПРОВАДЖЕННЯ ВІДКРИТИХ ДАНИХ

Першим кроком для впровадження політики відкритих даних на місцевому рівні може бути створення робочої групи, до складу якої варто залучити представників ІТ-департаменту, керівників інших департаментів, а також зацікавлених сторін. Бажано, щоб склад групи не перевищував 12 осіб. Робоча група повинна мати повноваження дорадчого органу, а її діяльність має бути закріплена у Плані дій із розвитку відкритих даних.

Завданнями робочої групи мають бути розробка плану дій із розвитку відкритих даних, загальний моніторинг реалізації завдань, напрацювання змін до нормативно-правової бази.



ПЛАН ДІЙ ІЗ РОЗВИТКУ ВІДКРИТИХ ДАНИХ

План дій із розвитку відкритих даних передбачає перелік завдань, часові межі та відповідальних за їх реалізацію осіб/підрозділів. Затвердження Плану дій дасть змогу структурувати й систематизувати роботу органу виконавчої влади, місцевого самоврядування.

План дій можна розробляти на основі принципів Міжнародної хартії відкритих даних і Дорожньої карти з розвитку відкритих даних в Україні. Варто зосередитись на таких основних завданнях:

1. Нормативне забезпечення (погодження нормативно-правової бази з нормами чинного законодавства, затвердження необхідних для налагодження процесу розпорядчих документів);
2. Організаційно-кадрове забезпечення (створення робочої групи, визначення/створення відповідального органу, посади чи відповідальних осіб у структурних підрозділах, державних підприємствах);
3. Фінансове забезпечення (передбачення коштів на розробку власного порталу відкритих даних, інтерфейсів прикладного програмування до баз даних, розробка баз даних, навчання, проведення хакатонів, розробка сервісів, продуктів на основі відкритих даних);
4. Аудит даних (організація та проведення аудиту, впровадження рекомендацій);
5. Контроль якості та доступності даних (моніторинг оприлюднених наборів даних);
6. Оприлюднення якісних даних (визначення переліку наборів для оприлюднення, пріоритетів);
7. Популяризація даних і співпраця з користувачами (організація та проведення зустрічей із зацікавленими сторонами, хакатонів);
8. Навчання (організація та проведення навчальних заходів, консультування).

Обов'язковою умовою ефективного впровадження Плану дій є чітке формулювання та реалістичність завдань, термінів і контроль за виконанням.



ПРИКЛАДИ:

- План дій із реалізації принципів Міжнародної хартії відкритих даних Львівської міської ради;
- План дій із реалізації принципів Міжнародної хартії відкритих даних Вінницької міської ради.

ВІДПОВІДАЛЬНІ ЗА ВІДКРИТІ ДАНІ

Відповідальною структурною одиницею за впровадження політики відкритих даних може бути ІТ-департамент (чи інший департамент, управління, відділ, сектор, який відповідає за інформаційно-комунікаційні технології, комп'ютерно-технічне забезпечення). Для цього потрібно внести зміни до Положення про департамент. У посадові інструкції працівників потрібно внести зміни або призначити/створити нову посаду – відповідального за відкриті дані.

Крім того, варто розглянути можливість створення окремого органу, відповідального за відкриті дані. У такому разі потрібно розробити Положення про орган, у якому визначити й закріпити його основні функції: впровадження, координація, контроль за відкритими даними, комунікація з користувачами, організація та проведення заходів.

До посадових інструкцій та/або положення про роботу департаменту, відповідального за відкриті дані в органі державної влади або місцевого самоврядування можна включити такі обов'язки:

1. Організація та проведення регулярного (не рідше ніж раз на рік) аудиту (інвентаризації) даних розпорядників публічної інформації відповідно до затвердженої методики;
2. Запровадження Реєстру даних розпорядників публічної інформації;
3. Підтримка та оновлення Реєстру даних розпорядників публічної інформації;
4. Контроль за виконанням розпорядниками публічної інформації нормативно-правових документів щодо відкритих даних;



5. Розробка змін до нормативно-правових документів щодо відкритих даних у разі потреби;
6. Розробка та контроль за виконанням Плану дій із розвитку відкритих даних;
7. Оптимізація роботи з даними, обміну даними між структурними підрозділами, державними, комунальними підприємствами, непідпорядкованими органами виконавчої влади, комерційними та громадськими організаціями;
8. Забезпечення внесення питань відкритих даних до розробки технічних рішень, які планують розробляти в органі виконавчої влади, місцевого самоврядування;
9. Забезпечення вивантаження даних із баз даних у форматі відкритих даних;
10. Забезпечення наповнення Єдиного державного порталу відкритих даних України, а за наявності також власного порталу відкритих даних або відповідного розділу власного сайту;
11. Забезпечення комунікації з користувачами відкритих даних та розпорядниками публічної інформації, надання їм консультацій;
12. Забезпечення регулярного оновлення наборів даних у форматі відкритих даних;
13. Своєчасне інформування керівництва про відхилення від термінів виконання робіт у разі наявних об'єктивних причин;
14. Налагодження, координація й контроль роботи департаментів щодо публікації даних у форматі відкритих даних на Єдиному державному порталі відкритих даних, а за наявності також на власному порталі відкритих даних, відповідному розділі сайту підпорядкованими підрозділами та підприємствами, установами, організаціями;
15. Регулярна (не рідше ніж раз на квартал) перевірка роботи розпорядників щодо публікації даних на Єдиному державному порталі відкритих даних, а за наявності також на власному порталі відкритих даних чи відповідному розділі сайту; перевірка якості оприлюднених даних;
16. Забезпечення комунікації з активістами, громадськими організаціями, науковим, IT та бізнес-середовищем щодо розвитку відкритих даних;
17. Організація та проведення громадських обговорень для визначення пріоритетних даних;
18. Організація та проведення регулярних (не рідше ніж раз на півроку) зустрічей із розпорядниками публічної інформації й непідпорядкованими органами виконавчої влади, іншими зацікавленими сторонами та власниками суспільно важливих даних щодо розвитку відкритих даних;



19. Організація та проведення навчання з питань відкритих даних для відповідальних за підготовку та оприлюднення даних (регулярність – за потреби);
20. Робота з розпорядниками та розробниками щодо використання даних для управління, ухвалення рішень, створення сервісів;
21. Підготовка звітів щодо публікації розпорядниками публічної інформації на Єдиному державному порталі відкритих даних, а за наявності також на власному порталі відкритих даних або відповідному розділі на сайті (щокварталу).

Відповідальний орган/особа за відкриті дані має право:

1. Вносити пропозиції щодо питань, які пов'язані з роботою у сфері відкритих даних, ухвалювати відповідні рішення;
2. Представляти орган влади з питань, які стосуються відкритих даних та їх використання;
3. Брати участь у нарадах та інших заходах, які пов'язані з інформаційно-комунікаційними технологіями й даними;
4. Отримувати необхідну для виконання своїх завдань і функцій інформацію, звіти від інших підпорядкованих органів виконавчої влади, місцевого самоврядування, структурних підрозділів, установ та організацій, державних підприємств;
5. Здійснювати аудит даних державних підприємств, які є розпорядниками публічної інформації.

У кожному структурному підрозділі розпорядчим документом потрібно визначити особу, яка оприлюднює (або надає) дані, набори даних. До її посадової інструкції потрібно додати такі обов'язки:

1. Підготовка даних до оприлюднення, оприлюднення (за потреби) відкритих даних на Єдиному державному порталі відкритих даних, а також на власному порталі відкритих даних або відповідному розділі сайту;
2. Робота з іншими структурними підрозділами, розробниками для покращення якості управління та створення сервісів;
3. Надання відомостей, звітів про роботу структурного підрозділу щодо відкритих даних до профільного структурного підрозділу.



ПОРЯДОК ОПРИЛЮДНЕННЯ ВІДКРИТИХ ДАНИХ

Порядок (положення, регламент) оприлюднення відкритих даних визначає роботу органу виконавчої влади або місцевого самоврядування щодо відкритих даних. Порядок може містити такі розділи:

1. Перелік наборів даних для оприлюднення зі вказанням відповідального структурного підрозділу, державного підприємства, частоти оновлення та форматів (може бути окремим документом);
2. Аудит даних (у розділі “Аудит даних” запропонована методика);
3. Ведення Реєстру даних (містить інформацію про всі дані, які є в органі виконавчої влади, місцевого самоврядування й оновлюється за повідомленням розпорядників);
4. Формати і стандарти управління даними, процедура оприлюднення даних на Єдиному державному порталі відкритих даних, а за наявності також на власному порталі, відповідному розділі сайту (у розділі “Організаційно-технічний аспект відкритих даних” запропоновано формати та стандарти);
5. Використання даних і комунікація з користувачами (у розділі “Використання даних” запропоновано дії).

Варто зазначити, що розпорядники публічної інформації звільняються від відповідальності за інтерпретацію даних, продукти та сервіси, які створені сторонніми користувачами на основі оприлюднених наборів даних.

ПРИКЛАДИ:

- Положення про збір та публікацію даних на Порталі відкритих даних Львова;
- Положення про відкриті дані Виконавчого комітету Дрогобицької міської ради та її виконавчих органів та впровадження порталу відкритих даних міста;
- Порядок про відкриті дані Чернівецької міської ради.



ПЕРЕЛІК НАБОРІВ ДАНИХ ДЛЯ ОПРИЛЮДНЕННЯ

Набори даних для першочергового оприлюднення визначені Постановою Кабінету Міністрів України №835 “Про затвердження Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних” (зі змінами). Структура й формат даних для оприлюднення визначається за результатами аудиту та затверджуються розпорядчим документом.

Окрім того, набори даних, які є пріоритетними для оприлюднення, але не передбачені постановою, включаються до Переліку за рекомендацією робочої групи. Проект Переліку наборів даних для оприлюднення готується відповідальним органом/особою за відкриті дані за результатами:

1. Аудиту даних;
2. Онлайн-опитування (з переліком даних для пріоритизації чи з відкритим запитанням);
3. Зустрічей, обговорень із зацікавленими сторонами, громадськістю;
4. Аналізу запитуваної громадянами та юридичними особами інформації в порядку доступу до публічної інформації, звернень щодо оприлюднення інформації;
5. Аналізу відвідуваності, завантаження, переглядів на офіційному сайті.

Перелік наборів даних для оприлюднення повинен містити назву набору даних, формат, термін, до якого набір повинен бути оприлюднений, частоту оновлення, відповідального за оприлюднення. Відповідальний орган/особа за відкриті дані здійснює контроль та нагляд за оприлюдненням наборів даних відповідно до Переліку наборів даних для оприлюднення.

ПРИКЛАДИ:

- Перелік наборів даних для оприлюднення на Порталі відкритих даних Чернівців;
- Реєстр наборів даних, які підлягають оприлюдненню у формі відкритих даних на офіційному веб-порталі Кременчуцької міської ради.



ПАМ'ЯТКА ДЛЯ САМОПЕРЕВІРКИ

1. Проаналізували національне законодавство та підзаконні акти, які є у сфері відкритих даних.

Так: Вітаємо! Ви все робите правильно.

Ні: Під час розробки власної нормативно-правової бази ви можете зробити помилки, які потім будуть оскаржені чи зроблять документи недієвими.

2. Проаналізували нормативно-правові документи, які регламентують доступ до публічної інформації, обмін інформацією між структурними підрозділами, зовнішніми користувачами.

Так: Вітаємо! Ви все робите правильно.

Ні: Нормативні акти можуть дублювати одне одного і збільшувати навантаження на працівників. Потрібно проаналізувати ті документи, які були затверджені й наразі діють, визначити, які з них втратять чинність після ухвалення нових нормативних документів, які потрібно включити до збірного документу тощо.

3. Визначили відповідальних структурних підрозділів/осіб за відкриті дані.

Так: Вітаємо! Ви все робите правильно.

Ні: Якщо не буде відповідальної структури чи особи в органі виконавчої влади, місцевого самоврядування, то процес впровадження і розвитку відкритих даних відбуватиметься хаотично, несистемно, керівництву буде важко отримати загальну картину щодо стану розвитку відкритих даних у підпорядкованих йому структурних підрозділах, підвідомчих установах, державних підприємствах тощо. Відсутність відповідальних за відкриті дані може негативно вплинути на якість оприлюднених даних на Єдиному державному порталі відкритих даних України, власному порталі, відповідному розділі на сайті, а також на комунікації з користувачами, розробниками тощо.



4. Визначили Перелік наборів даних для оприлюднення, відповідальних та частоти оновлення.

Так: Вітаємо! Ви все робите правильно.

Ні: Це може призвести до того, що структурні підрозділи будуть вільно тлумачити Постанову №835 (зі змінами), а оприлюднення даних відбуватиметься несистематично, набори будуть мати погану якість.

5. Розробили Порядок оприлюднення відкритих даних.

Так: Вітаємо! Ви все робите правильно.

Ні: Нормативна база з питань відкритих даних буде розкидана в інших нормативно-правових документах. Єдиного зведення правил роботи з відкритими даними, форматів та стандартів, принципів роботи не буде. Це призведе до того, що впровадження політики відкритих даних відбуватиметься без чітких правил і розуміння, тобто до низької якості.



АУДИТ ДАНИХ

У цьому розділі ви дізнаєтеся більше про мету, етапи й результати аудиту, а також про повноваження аудитора й реєстр даних.

Аудит даних – це дослідження наявності, стану, форматів, процесів управління й використання даних, а також вироблення на основі отриманої інформації рекомендацій щодо покращення процесів роботи з даними, максимізації їх використання та розкриття потенціалу.

Проведення аудиту даних рекомендоване для виконання Постанови Кабінету Міністрів України №835 “Про затвердження Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних” зі [змiнами](#) від 20 грудня 2017 року, яка визначає перелік наборів, що обов’язкові до оприлюднення. Аудит дає змогу розпоряднику дослідити й зрозуміти, у якому стані наразі дані і що потрібно зробити, щоб налагодити процес оприлюднення необхідних наборів у форматі відкритих даних.

Аудит допомагає також виявити дублювання зусиль розпорядників щодо збору та оприлюднення даних, виявляє проблемні ділянки, які потребують додаткової роботи, що дає змогу краще розподілити ресурси.

Для проведення аудиту даних потрібні:

Політична воля – керівництво органу виконавчої влади, місцевого самоврядування підтримує аудит даних, надаючи повноваження аудитору та співпрацюючи з ним під час розробки та впровадження рекомендацій. Це може відбуватись за допомогою ухвалення відповідних розпорядчих документів, які регулюють проведення аудиту.



Аудитор – особа або група осіб, які мають необхідні навички роботи з даними, а також повноваження та доступ до даних розпорядників. Аудит даних може проводитись як працівниками органів виконавчої влади, місцевого самоврядування (бажано, щоб це були працівники ІТ-управлінь, департаментів, відділів, секторів, правники), так і зовнішніми експертами.

Співпраця та виконання – розпорядники повинні допомагати аудитору виявити дані та зрозуміти їхній стан і потенціал, а також спільно розробляти рекомендації та втілювати їх на практиці.

Періодичність проведення аудиту визначається органом центральної влади, місцевого самоврядування, згідно з Порядком оприлюднення наборів даних, Планом дій з розвитку відкритих даних).

ОРГАНІЗАЦІЯ АУДИТУ ДАНИХ

Для організації аудиту потрібно визначити:

1. Відповідальний орган/особа за проведення аудиту;
2. Часові рамки проведення аудиту;
3. Розпорядники, у яких потрібно здійснити аудит (визначаються відповідно до переліку наборів даних, визначених Постановою Кабінету Міністрів України №835 “Про затвердження Положення про набори даних, які підлягають оприлюдненню у формі відкритих даних” (зі змінами), а також зважаючи на суспільний запит).

Робоча група пропонує вид аудиту даних, який буде використовуватись органом виконавчої влади, місцевого самоврядування.



ВИДИ АУДИТУ ДАНИХ

1. **Віддалений аудит.** Інформація про стан і якість даних збирається від розпорядників даних через відповідні форми. Розпорядники заповнюють форму визначеної структури (здебільшого вона містить запитання про те, які дані має розпорядник, у якому форматі, частоту оновлення тощо) і разом із прикладами наборів надсилають її аудитору для аналізу. На її основі аудитор робить аналіз даних і надає свої рекомендації. Такий формат аудиту не потребує особистого спілкування аудитора з працівниками, роботи з базами даних та опрацювання зразків наборів даних. Цей варіант аудиту потребує менше часу аудитора, однак може не надати повної картини стану даних.
2. **Контактний аудит.** Аудитор попередньо досліджує роботу органу, де відбувається аудит, опитує працівників, вивчає бази даних і технології в розпорядника, отримує зразки даних, самостійно заносить їх у форму з документування даних, аналізує й розробляє рекомендації. Цей формат аудиту потребує більших затрат часу аудитора та працівників органу, де відбувається аудит. Він дає більш докладну й точну інформацію.
3. **Комбінований аудит.** Аудитор надає розпорядникам форми для заповнення і працює з ними безпосередньо. Після отримання форм від розпорядників, аудитор опрацьовує їх разом з отриманими під час контактного аудиту даними та розробляє на їх основі рекомендації. Варто, щоб віддалений і контактний аудити були розділені в часі. Контактна частина аудиту має проводитися після отримання й опрацювання таблиць із даними. Комбінований аудит дає максимально повну картину, однак потребує більше часу аудитора та розпорядників.

Проведення аудиту даних визначається розпорядчим документом (рішенням, розпорядженням, наказом), у якому повинні бути зазначені такі питання:

1. Предмет аудиту даних (що саме підлягає аудиту);
2. Терміни проведення всіх етапів аудиту;
3. Графік із вказанням органу, де проводиться аудит, і його контактної особи (у разі контактного аудиту);



4. Перелік структурних підрозділів, державних, комунальних підприємств (органів), у яких передбачено аудит даних;
5. Перелік осіб, які здійснюватимуть аудит даних (аудиторів);
6. Повноваження аудиторів:
 - отримувати допомогу й підтримку від контактної особи органу, у кому проводиться аудит;
 - мати доступ до документів, баз даних, які є в розпорядженні органу, де відбувається аудит, переліку інформації, технологій і систем;
 - мати можливість отримати зразки даних для опрацювання та аналізу;
 - мати можливість провести інтерв'ю зі всіма працівниками, управлінцями, IT-фахівцями органу, де відбувається аудит;
6. Обов'язки керівництва і працівників органів, де відбувається аудит;
7. Процедура розгляду та впровадження рекомендацій за результатами аудиту даних.

Аудитор повинен мати базові знання щодо інформаційно-комунікаційних технологій, процесів управління даних, роботи з даними. Необхідним є знання чинного законодавства та ознайомлення з нормативно-правовою базою органів виконавчої влади, місцевого самоврядування.

Аудит даних органу виконавчої влади, місцевого самоврядування може тривати 2-6 місяців залежно від кількості органів, де відбувається аудит.



ЕТАПИ ПРОВЕДЕННЯ АУДИТУ ДАНИХ

ДОКУМЕНТУВАННЯ ДАНИХ

Цей етап передбачає роботу з розпорядником (віддалену чи персональну). Залежно від того, який формат буде обрано, аудитор документує дані самостійно або ні.

Документування даних охоплює дослідження таких характеристик даних:

Identifier	Унікальний ідентифікатор, визначений аудитором для набору даних
Title	Назва набору даних
Creator	Розпорядник, відповідальний за набір даних
Contact	Засоби зв'язку з відповідальною за набір особою
Description	Короткий опис набору даних, включно з просторовими, часовими та предметними показниками
DateCreated	Дата, коли дані були зібрані чи створені
Purpose	Підстава (нормативний документ) для створення цієї інформації
Source	Джерело інформації в цьому наборі даних (якщо є)
Receiver	Отримувачі інформації (наприклад, вищий орган влади, якому розпорядник передає ці дані)
Updating frequency	Частота оновлення набору даних
Usage frequency	Частота використання даних
Public interest	Частота звернень щодо інформації з цього набору даних
Relation	Опис зв'язків з іншими наборами даних, базами даних
Link	Якщо набір оприлюднено, то надати посилання
Format	Кінцевий формат даних (на цей момент)
Type	Кінцевий тип даних, відповідно до переліку, наведеному в Постанові №835 (зі змінами)



Structure	Деталі структури набору даних (наприклад, назви колонок у таблиці)
Documentation	Документація, яка є для цього набору даних
Dictionary	Довідники, які доступні для набору даних
File size	Розмір набору даних (Мб, Гб)
Quality	Стан і якість набору даних
Database	Назва бази даних, де міститься інформація (якщо таке є)
Personal Data	Персональна інформація, яка є в наборі (якщо така є)
Back-up	Кількість копій набору даних, які наразі зберігаються, частота створення резервних копій і процедур архівування
Management to date	Права редагування, безпека, підтримка, оновлення
Requirements	Опис будь-якого спеціалізованого обладнання чи програмного забезпечення, необхідного для створення чи утримання набору даних
Competences	Компетенції, які потрібні відповідальним працівникам для роботи з даними

Отримані дані заносяться в таблицю http://bit.ly/audit_data_structure.

У разі проведення віддаленого аудиту аудитор надає вказану таблицю контактній особі органу, у якому відбувається аудит, і пояснює, як її заповнювати. Якщо в органі, де відбувається аудит, є кілька структурних підрозділів, то контактна особа зобов'язана надати таблицю для заповнення кожному підрозділові, а потім зібрати отримані таблиці й передати на опрацювання аудитору.

Під час контактного аудиту даних аудитор спілкується з працівниками органу та збирає інформацію про управління даними відповідно до затвердженого графіку та таблиці. Окрім інформації, яка є в таблиці з документування даних, аудитор зважає на інформацію, яку він отримав під час підготовки до аудиту. У робочий документ заносяться власні спостереження.

За результатом аудиту формується аналітичний звіт, який має містити детальний опис реального стану предмету аудиту та рекомендації щодо покращення. На основі такого звіту орган, у



якому відбувався аудит, має усунути виявлені недоліки відповідно до наданих рекомендацій. Контроль за виконанням рекомендацій має здійснювати керівництво органу.

На основі результатів аудиту має бути сформовано та затверджено Реєстр публічної інформації у формі відкритих даних органу виконавчої влади / місцевого самоврядування.

РОЗРОБКА ТА ПРЕДСТАВЛЕННЯ РЕКОМЕНДАЦІЙ НА ОСНОВІ АУДИТУ

За результатами документування даних аудитор розробляє рекомендації. Основою для їх написання є робочий документ і таблиці з характеристиками даних органу, де відбувався аудит.

Рекомендації за результатами аудиту повинні бути представлені у вигляді звіту й надані на розгляд робочій групі зі впровадження політики відкритих даних.

Рекомендації повинні мати таку структуру:

1. Загальна характеристика органу виконавчої влади, місцевого самоврядування (аудитор дає огляд нормативної бази з питань відкритих даних, стану управління даними загалом, ґрунтуючись на отриманій інформації під час підготовки та проведення аудиту);
2. Загальні рекомендації органу виконавчої влади, місцевого самоврядування (поради щодо того, як забезпечити якісне оприлюднення та використання даних);
3. Рекомендації для кожного конкретного органу, де відбувався аудит, з описом ситуації на момент аудиту й порадами щодо того, як оприлюднювати набори даних, скільки ресурсів потребує публікація наборів даних тощо;
4. Додатки (аудитор також може розробити типову структуру для наборів даних, опис набору даних та запропонувати частоту оновлення).

Рекомендації, які складаються з описової частини та зведеної таблиці за результатами документування даних, надаються відповідальному органу/особі за відкриті дані для опрацювання.



Аудитор спільно з відповідальним органом/особою за відкриті дані готують презентацію рекомендацій за результатами аудиту. Звіт щодо впровадження рекомендацій публікується на сайті органу виконавчої влади, місцевого самоврядування відповідальним органом/особою. Аналітичний звіт, що є результатом аудиту також публікується на сайті органу виконавчої влади, місцевого самоврядування.

На основі рекомендацій відкритих даних члени робочої групи розробляють проект розпорядчого документу щодо першочергових дій для їх впровадження, які є обов'язковими до виконання розпорядниками.

Для кожного органу виконавчої влади чи місцевого самоврядування ведеться реєстр наборів даних. Його адміністрування й оновлення забезпечує відповідальний орган/особа за відкриті дані на основі даних, отриманих під час аудиту, а згодом від розпорядників. Зміни до реєстру вносяться у міру появи нових наборів даних чи внесення змін до набору даних, а також за результатами чергового аудиту.

Розпорядники публічної інформації зобов'язані повідомляти відповідальний орган/особу за відкриті дані про новий набір чи зміни до вже наявного набору. Те саме стосується змін структури, формату набору даних, частоти оновлення, розпорядника даних. Відповідальна особа за відкриті дані вносить зміни до Реєстру даних у міру надходження відомостей від розпорядника.

Реєстр даних публікується на сайті та Єдиному державному порталі відкритих даних у форматі відкритих даних. Реєстр використовують розпорядники для оперативного пошуку та використання потрібних даних, для комунікацій із громадськістю щодо визначення пріоритетних даних до публікації.



Структура Реєстру наборів даних має бути така:

Identifier	Унікальний ідентифікатор набору даних
Title	Назва набору даних
Description	Короткий опис
DateCreated	Дата створення набору даних
Format	Формат набору даних
Type	Тип набору даних відповідно до переліку, наведеного в Постанові № 835 (зі змінами)
File size	Розмір набору даних
Update frequency	Частота оновлення набору даних
Link	Посилання на набір даних, якщо він оприлюднений
Creator	Назва розпорядника, відповідального за набір даних
Contact	Засоби зв'язку з відповідальним



ПАМ'ЯТКА ДЛЯ САМОПЕРЕВІРКИ

1. Затвердили порядок проведення аудиту даних і повноваження аудитора.

Так: Вітаємо! Ви все робите правильно.

Ні: Аудитор не отримає доступу до всієї інформації, даних, баз даних, які має в розпорядженні орган, де проходить аудит, а також не матиме повноважень проводити опитування працівників. Результати аудиту не матимуть ніякого впливу на процеси управління й оприлюднення даних.

2. Роз'яснили цілі та завдання аудиту даних для розпорядників, повноважень аудитора й обов'язків розпорядників, де відбуватиметься аудит.

Так: Вітаємо! Ви все робите правильно.

Ні: Можна очікувати, що розпорядники будуть поверхово ставитись до запитань аудитора з остраху, що їм дадуть ще більше роботи, не будуть вносити дані, які вони мають у розпорядженні тощо.

3. Представили результати аудиту даних органів виконавчої влади, місцевого самоврядування й ухвалили рішення щодо впровадження його рекомендацій.

Так: Вітаємо! Ви все робите правильно.

Ні: У такому разі аудит даних може залишитися лише у вигляді рекомендацій та не мати жодного впливу на покращення якості управління даними, їх використання.

4. Створили й підтримуєте Реєстр наборів даних.

Так: Вітаємо! Ви все робите правильно.

Ні: Якщо немає систематизованої інформація про те, які, де і в якому стані є дані, то процес пошуку даних самими розпорядниками, визначення пріоритетних наборів даних для



оприлюднення потребуватиме більше часу. Також є небезпека того, що наступний аудит буде ускладнений, якість управління даними залишиться на тому ж рівні, що і до аудиту.

5. Перевірили та прозвітували, чи впроваджуються рекомендації аудиту.

Так: Вітаємо! Ви все робите правильно.

Ні: Брак контролю та нагляду призведе до того, що розпорядники не будуть виконувати рекомендації.



ПІДГОТОВКА ДАНИХ ДО ПУБЛІКАЦІЇ

Публікація даних потребує ретельного планування.

Цей розділ має на меті дати розпорядникам базовий алгоритм підготовки наборів даних до оприлюднення та налагодження процесу публікації й оновлення даних на Єдиному державному порталі відкритих даних.

Перш ніж публікувати дані, потрібно провести їх ревізію. Якщо ви маєте оприлюднити набори даних, визначені Постановою Кабінету Міністрів України №835 (зі змінами), візьміть цей перелік наборів даних і складіть таблицю, у якій зазначте такі пункти:

- Назва набору даних;
- Розпорядник набору даних;
- Відповідальна особа за ведення набору даних;
- Формат ведення і збереження набору даних (наприклад, автоматизована інформаційна система, файл у форматі XLS(X), один або кілька файлів у форматі DOC(X));
- Тип даних у наборі (наприклад, текстові чи структуровані);
- Частота оновлення даних;
- Наявність/відсутність у наборі персональних даних службової інформації або іншої інформації, що не підлягає оприлюдненню.

Якщо ви вже робили аудит наборів даних, це завдання не забере у вас багато часу. У іншому разі вона стане гарним починком для проведення повноцінного аудиту даних.

На цьому ж етапі вам важливо для себе з'ясувати низку питань.



Які дані потенційно мають найбільшу цінність для користувачів? Традиційно найбільш цінними для користувачів є структуровані дані у машиночитаних форматах. Саме вони зазвичай стають основою для аналітичних матеріалів, журналістських розслідувань, сервісів і застосунків. Якщо ви маєте такі набори даних, варто приділити їм першочергову увагу.

Чи відповідає формат типам даних, які в них містяться? Наприклад, ви маєте автоматизовану інформаційну систему, у якій збираються та зберігаються структуровані (табличні) дані. За таких умов із великою вірогідністю ви зможете без проблем публікувати структуровані дані у машиночитаних форматах. Або ви маєте реєстр у табличному вигляді, який ведеться у файлі формату DOC(X). У такому разі формат файлу не відповідає типу даних, що в ньому містяться, а отже вам доведеться трансформувати ці дані перед публікацією.

Який формат публікації набору даних є оптимальним? Постановою Кабінету Міністрів України №835 (зі змінами) передбачені різні формати для різних типів даних. Якщо у вашому наборі даних містяться структуровані (наприклад, табличні) дані, тоді їх треба оприлюднювати у форматі CSV. Якщо у наборі є структуровані та ієрархічні дані, для них підійдуть формати JSON і XML. Для текстових даних потрібно використовувати формат RTF. Якщо набір даних оновлюється щодня, то, можливо, для нього варто розробити інтерфейс прикладного програмування (API). У будь-якому разі вам потрібно визначити, який формат є оптимальним для кожного конкретного набору даних. А якщо формат ведення набору даних не збігається з оптимальним для нього форматом публікації, то потрібно ще й визначити процедуру трансформації даних.

Якою має бути частота оновлення даних? Згідно з Постановою №835 (зі змінами), розпорядники можуть самостійно визначати частоту оновлення оприлюднених даних. Частота може варіюватись від “відразу після внесення змін” до “щороку”. У кожному випадку варто зважати на те, як часто оновлюються дані в самому наборі (реєстрі, базі даних). Додатковим критерієм може бути наявність запиту на оновлення даних від потенційних користувачів. Наприклад, якщо робота певного сервісу, застосунку чи аналітичного інструменту потребує щотижневого оновлення даних, є сенс оновлювати дані кожного тижня.



Чи потрібно буде деперсоналізувати дані перед публікацією? Персональними даними, відповідно до закону “Про захист персональних даних”, є відомості чи сукупність відомостей про фізичну особу, яка ідентифікована або може бути конкретно ідентифікована. Наприклад, це може бути прізвище, ім’я, по батькові, дата й місце народження, місце проживання, паспортні дані й ідентифікаційний код тощо. Якщо у наборі даних, який ви маєте оприлюднити, містяться персональні дані, потрібно визначити процедуру деперсоналізації даних. Це питання також є важливим під час визначення періодичності, з якою набір даних буде оновлюватися.

Скільки часу потребує підготовка даних до публікації? Відповіді на попередні питання допоможуть вам краще зрозуміти обсяг робіт із підготовки наборів даних до публікації, визначити свої організаційні та технічні потреби, а також призначити відповідальних осіб і терміни виконання поставлених завдань.

ВИЗНАЧЕННЯ ФОРМАТУ ПУБЛІКАЦІЇ ДАНИХ

Одним із найбільш поширених питань, що виникають у розпорядників даних, є питання вибору файлового формату для публікації наборів даних. Цей розділ має на меті пояснити, як обрати файловий формат для найбільш поширених типів даних.

Постановою Кабінету Міністрів України №835 (зі змінами) передбачено публікацію декількох типів даних, як-от текстових, структурованих, геопросторових, графічних, відео- та аудіоданих, розроблених із використанням програми Macromedia Flash, архівів даних.

Важливо те, що для кожного типу даних Постановою встановлені відповідні файлові формати.

Тип даних	Формат
Текстові дані	TXT, RTF, ODT, DOC(X), PDF (з текстовим змістом, нескановане зображення), (X)HTML
Структуровані дані	RDF, XML, JSON, CSV, XLS(X), ODS, YAML



Тип даних	Формат
Геопросторові дані	GeoTIFF, SHP, DMF, MID/MIF, DXF, XML, GeoJSON, GPX, LOC, ARINC, AIXM
Графічні дані	GIF, TIFF, JPG (JPEG), PNG
Відеодані	MPEG, MKV, AVI, FLV, MKS, MK3D
Аудіодані	MP3, WAV, MKA
Дані, розроблені з використанням програми Macromedia Flash	SWF, FLV
Архів даних	ZIP, 7z, Gzip, Bzip2

Під час визначення формату для оприлюднення набору даних необхідно зважати саме на відповідність типу даних файловому формату. Найбільш поширеною помилкою, яка виникає під час публікації наборів даних, є невідповідність файлового формату типу даних, що у ньому міститься. Зокрема, некоректною є публікація таблиць (структурованих даних) у форматах DOC(X) чи PDF, призначених для текстових даних, або у форматах JPG чи PNG, призначених для графічних даних.

Крім того, варто мати на увазі, що під час створення нових наборів даних перевага має надаватись відкритим файловим форматам, тобто таким, що не залежать від платформи та доступні без обмежень, які можуть перешкодити їх повторному використанню. До відкритих форматів, зокрема, належать формати, як-от ODT, HTML, RDF, XML, JSON, CSV, ODS, YAML.

ТЕКСТОВІ ДАНІ

Якщо більшість місця у вашому наборі займає простий текст – ви маєте справу з текстовими даними. Прикладом текстових даних можуть бути нормативно-правові акти, положення, рішення чи розпорядження.



Публікувати текстові дані потрібно передусім у відкритих форматах RTF та ODT. Дозволяються також формати DOC(X) і PDF (з несканованим зображенням).

Категорично не підходять для текстових даних формати JPG, JPEG, PNG, GIF, TIFF, а також PDF зі сканованим зображенням. Публікація текстових даних у цих форматах унеможлиблює їх обробку автоматизованими засобами.

Якщо ви плануєте публікувати багато типових наборів даних у текстових форматах, наприклад, рішень міської ради, є сенс додатково створити таблицьку у форматі CSV, у якій буде подано перелік цих рішень. Наприклад, зі зазначенням дати ухвалення, ідентифікаційного номеру, заголовку чи опису одним реченням та назви файлу, у якому міститься повний текст рішення чи посилання на нього.

Подібний формат публікації рішень практикує Державна судова адміністрація.

У такому разі для кожного рішення потрібно не створювати окремий набір даних на Єдиному державному порталі відкритих даних, а додавати їх у один із загальною назвою “Рішення...” або ж “Розпорядження...”.

СТРУКТУРОВАНІ ДАНІ

Якщо у вашому наборі даних є таблиці, значить, ви маєте справу зі структурованими даними. Прикладами структурованих даних можуть бути різноманітні реєстри, переліки, плани та звіти.

Публікувати структуровані дані потрібно насамперед у відкритих форматах CSV та ODS. Дозволяється також формат XLS(X).

Якщо ваші дані зберігаються в інформаційній системі чи базі даних, яка дає змогу вивантажувати дані у форматах XML чи JSON, варто використовувати їх для публікації набору



даних. Втім, зважайте на те, що найкраще формати XML і JSON підходять для ієрархічних даних. Якщо ваші дані не є ієрархічними за природою, для їх публікації буде достатньо формату CSV.

Для публікації структурованих (табличних) даних категорично не підходять формати DOC(X), RTF, PDF, JPG, JPEG, TIFF, PNG.

ГЕОПРОСТОРОВІ ДАНІ

Якщо набір даних містить інформацію про розташування певних об'єктів із зазначенням широти й довготи, або опис меж певних територій із використанням полігонів, ви маєте справу з геопросторовими даними.

Прикладом геопросторових даних можуть бути генеральні плани населених пунктів, схеми планування територій і плани зонування територій, межі виборчих округів та діляниць, відомості з Держгеокадастру, маршрути й дані про місцезнаходження громадського транспорту і т. д.

Геопросторові дані мають публікуватися передусім у відкритих форматах GeoJSON, SHP, GPX, GeoTIFF. Також можна використовувати формати DMF, MID/MIF та інші.

Для публікації геопросторових даних за поодинокими винятками не підходять формати PNG, JPG чи JPEG, призначені для графічних даних.

Деякі набори даних, наприклад, “Інформація про рекламні засоби”, “Дані про розміщення громадських вбиралень” або “Дані про місце розміщення зупинок”, можуть бути збагачені геоданими, а саме географічними координатами, що позначають точне розташування об'єкта. Однак такі набори даних цілком можна публікувати у звичайних табличних форматах CSV чи ODS.



ГРАФІЧНІ ДАНІ

Якщо набір даних є зображенням, ви маєте справу зі графічними даними. Прикладом таких даних можуть бути схеми планування територій і генеральні плани міст, але лише за умови, що вони відсутні у форматі геопросторових даних.

Якщо ви змушені публікувати геопросторові дані у файлових форматах, призначених для графічних даних (скажімо, за браком доступу до первинних геоданих), варто зазначити в метаданих географічні координати чотирьох крайніх точок на зображенні, щоб уможливити його прив'язку до карти.

Графічні дані мають публікуватись у відкритих форматах PNG, JPG чи JPEG.

АРХІВИ

Якщо ваш набір даних міститься у файлі великого розміру, або ви публікуєте багато типових файлів, що є частиною одного набору даних, є сенс використовувати для публікації архіви даних. Вони допомагають зменшити розмір набору даних і завантажити велику кількість типових файлів за один раз.

Для публікації архівів даних насамперед треба використовувати відкриті формати ZIP та 7z. Не варто використовувати для публікації архівів даних формат RAR, який є пропрієтарним.

Якщо ви використовуватимете архів даних, переконайтесь, що ви надаєте користувачам усю необхідну інформацію про вміст архіву (формат файлів, опис структури даних) у паспорті набору даних.



ІНТЕРФЕЙС ПРИКЛАДНОГО ПРОГРАМУВАННЯ (API)

Якщо ваш набір даних містить великий обсяг інформації й часто оновлюється, доступ до нього варто запровадити за допомогою інтерфейсу прикладного програмування.

Інтерфейс прикладного програмування містить набір функцій, що дає змогу отримати доступ до набору даних або певних його частин, створювати запити з різними параметрами. Він має забезпечувати можливість автоматизованого доступу до набору даних у цілодобовому режимі без вихідних.

Прикладами надання доступу до даних за допомогою інтерфейсу прикладного програмування можуть бути [Єдиний державний реєстр декларацій](#), [Єдиний веб-портал використання публічних коштів](#), а також [набори даних Національного банку України](#).

У разі запровадження інтерфейсу прикладного програмування для доступу до набору даних, треба також надати детальні інструкції щодо його використання, приклади запитів зі всіма можливими параметрами та приклади відповідей на ці запити.

СТРУКТУРУВАННЯ Й ОЧИЩЕННЯ ДАНИХ

Найбільш цінними для користувачів є саме структуровані та машиночитані дані. Однак із цим типом даних традиційно виникає найбільше проблем у розпорядників даних. Цей розділ має на меті прояснити вимоги до оформлення та публікації структурованих даних.

Відповідно до Постанови №835 (зі змінами), машиночитаним вважається формат даних, який дає змогу інформаційним системам ідентифікувати, розпізнавати, перетворювати й отримувати конкретні дані без участі людини.



Інформаційні системи й системи керування базами даних дають змогу вивантажувати набори даних у структурованих і машиночитаних форматах CSV, JSON чи XML. Проблеми виникають у тих випадках, коли набори даних ведуться у простих файлах на зразок XLS(X) чи DOC(X). Зазвичай подібні набори даних необхідно спеціально готувати до публікації, тобто приводити їх до правильної табличної структури.

ПРАВИЛА СТРУКТУРУВАННЯ ТАБЛИЧНИХ ДАНИХ

Ознаки правильної табличної структури:

- Усі змінні записані у стовпчиках;
- Усі спостереження записані у рядках;
- У таблиці немає об'єднаних комірок, один запис займає лише одну комірку;
- У таблиці немає об'єднаних записів, одна комірка містить лише один запис.

Наприклад, таблиця 1 має правильну структуру, оскільки в ній змінні (region, year, population) записані у стовпчиках, а спостереження (Київ, 2015, 2888.0 тощо) записані в рядках. Кожне спостереження займає лише одну комірку, а одна комірка містить лише одне спостереження.

Таблиця 1.

region	year	population
Київ	2015	2888.0
Київ	2016	2906.6
Київ	2017	2925.8

Таблиця 2 має неправильну структуру, оскільки в ній значення змінної (year) записані як назви змінних.



Таблиця 2.

region	2015	2016	2017
Київ	2888.0	2906.6	2925.8

Таблиця 3 теж має неправильну структуру, оскільки в ній один запис – “Київ” – займає аж три комірки.

Таблиця 3.

region	year	population
Київ	2015	2888.0
	2016	2906.6
	2017	2925.8

Аналогічно, неправильним є спосіб запису даних у таблиці 4, оскільки в ньому один запис (назва змінної “population”) займає дві комірки, а в значеннях цієї змінної змішано текст числа.

Таблиця 4.

region	year	population	
		available	permanent
Київ	2015	2888.0	2846.7
Київ	2016	2906.6	2865.3
Київ	2017	2925.8	2884.5



ЧИСТОТА ДАНИХ

Однак самої лише правильної структури недостатньо для коректної обробки даних автоматичними засобами. Дані мають бути чистими.

Чистими вважаються дані, у яких:

- Немає помилок чи одруківок, зайвих символів або пропусків;
- Для запису назв, дат, чисел вживається уніфікований формат;
- Для позначення відсутніх записів використовується NA, а не "0", "-" чи інші аналоги;
- У межах однієї змінної вживається лише один тип даних, тобто в одному стовпчику не змішуються, наприклад, числові й текстові значення.

У таблиці 5 наведено найпростіший зразок нечистих даних на прикладі вже знайомого нам набору. У змінній "region" присутні три різних способи запису назви міста Київ. У змінній "year" – два різних варіанти запису року та змішування текстових і числових значень. У змінній "population" – три різних варіанти запису числа, один з яких змішує текстові та числові значення.

Таблиця 5.

region	year	population
Київ	2015	2888.0
Києв	2016	2906.6
Кийів	'17	2,9258 млн

КОРИСНІ ІНСТРУМЕНТИ ПЕРЕВІРКИ ДАНИХ ПЕРЕД ПУБЛІКАЦІЄЮ

Для виявлення проблем у даних ви можете використовувати безкоштовну програму [Dataproofer](#). Вона застосовує до кожного набору даних 12 різних тестів, зокрема перевіряє наявність пропущених даних і дублікатів, підозрілі символи чи значення тощо.



Для очищення даних перед публікацією ви можете скористатись безкоштовною програмою [OpenRefine](#). Вона дає змогу швидко виявляти різні варіанти написання назв і приводити їх до однієї форми, видаляти дублікати та зайві символи, змінювати структуру даних, заповнювати порожні комірки тощо.

ПАМ'ЯТКА ДЛЯ САМОПЕРЕВІРКИ

Погано структуровані та “брудні” дані не є машиночитаними. Їх неможливо обробляти автоматизованими засобами без участі людини. Добре структуровані й чисті дані можна швидко аналізувати, візуалізувати, використовувати в сервісах чи дослідженнях. Тож перед публікацією табличних даних переконайтеся, що:

- У таблиці немає порожніх рядків на початку або всередині;
- У першому рядку записані назви змінних латинкою (для назв змінних бажано використовувати загальноприйняті словники на зразок [schema.org](#)). У рядках із другого й до останнього записані спостереження;
- У таблиці не використовується форматування (стиль, колір чи розмір шрифтів, колір заповнення комірок);
- У межах однієї змінної вживається один тип даних. Тобто, якщо в певному стовпчику записані числові дані, у ньому не має бути текстових даних;
- У межах однієї змінної використовується один формат запису дат, назв, чисел. Для роздільника десяткових значень у числах вживається крапка (наприклад, “10.1”, а не “10,1”). Для запису дат використовується формат РРРР-ММ-ДД (наприклад, 2017-12-31 для позначення 31 грудня 2017 року);
- Для позначення пропущених чи відсутніх значень використовується “NA” (а не 0, -, -99999 чи будь-яке інше значення);
- У даних нема дублікатів;
- Файл збережено з кодуванням UTF-8 у форматі CSV. Назва файлу записана латинкою.



ЗНЕОСОБЛЕННЯ ДАНИХ

Поняття персональних даних в Україні визначається Законом України “[Про захист персональних даних](#)”. Відповідно до цього закону, персональними даними можна вважати “відомості чи сукупність відомостей про фізичну особу, яка ідентифікована або може бути конкретно ідентифікована”.

До таких відомостей, наприклад, може належати прізвище, ім’я, по батькові, дата і місце народження, місце проживання, номер телефону, паспортні дані та ідентифікаційний код, національність, освіта, сімейний стан, релігійні та світоглядні переконання, стан здоров’я, матеріальний стан, дані про особисті майнові та немайнові відносини цієї особи з іншими особами.

Така інформація про фізичну особу та членів її сім’ї є конфіденційною і може поширюватись тільки за їхньою згодою, крім випадків, визначених законом, і лише в інтересах національної безпеки, економічного добробуту та захисту прав людини.

Не є конфіденційною інформацією персональні дані щодо здійснення посадових або службових повноважень особою, яка займає посаду, пов’язану з виконанням функцій держави або органів місцевого самоврядування.

До інформації з обмеженим доступом не належать персональні дані, зазначені в декларації про майно, доходи, витрати та зобов’язання фінансового характеру (окрім відомостей, визначених Законом України “Про засади запобігання і протидії корупції”).

Також не належить до інформації з обмеженим доступом інформація, що стосується отримання в будь-якій формі фізичною особою бюджетних коштів, державного чи комунального майна, крім випадків, передбачених статтею 6 Закону України “Про доступ до публічної інформації”.

Отже, перед публікацією наборів даних, треба видалити з них персональні дані, які належать до конфіденційної інформації або до інформації з обмеженим доступом.



ОФОРМЛЕННЯ НАБОРУ ДАНИХ

Оформлення набору даних під час публікації на Єдиному державному порталі відкритих даних – назва, опис і ключові слова, що вживаються для його характеристики, – не менш важливим, ніж підготовка самих даних, їх структурування й очищення. Якісно оформлений набір даних легко знайти. Його зміст просто зрозуміти з опису та метаданих. Цей розділ містить рекомендації щодо найважливіших елементів оформлення наборів даних під час публікації.

НАЗВА НАБОРУ ДАНИХ

Назва є критично важливою для пошуку та ідентифікації набору даних на Єдиному державному порталі відкритих даних. Коректна назва дає змогу швидко знайти й ідентифікувати набір даних. Некоректна назва ускладнює знаходження набору даних і зменшує шанси на його використання.

Якщо ви публікуєте набір даних, передбачений Постановою Кабінету Міністрів України №835 (зі змінами), використовуйте назву набору даних, зазначену в Постанові. Якщо ж ви публікуєте набір даних, який не передбачений Постановою, використовуйте його офіційну назву, зафіксовану в нормативно-правових актах, або її скорочену версію.

Не використовуйте без потреби в назві набору даних назву розпорядника, місто чи регіон його розташування, дати, номери чи будь-які інші ідентифікатори. Ці відомості безумовно важливі для ідентифікації набору даних, але їх варто зазначати в інших місцях, наприклад, у паспорті набору даних.

Ідеальна назва має бути коротка й загальна. З неї має бути відразу зрозуміло, про що йдеться в наборі даних.

Приклад некоректної назви набору даних: “Річний план закупівель на 2018 рік Українського центру оцінювання якості освіти (зі змінами від 07.02.2018 року)”.



Приклад ліпшої назви набору даних: “Річний план закупівель на 2018 рік”.

Приклад коректної назви набору даних: “Річні плани закупівель” – за умови, що всі річні плани закупівель конкретного розпорядника даних публікуються в одному наборі даних.

ОПИС НАБОРУ ДАНИХ

Опис потрібен для того, щоб користувачі мали можливість зрозуміти, про що йдеться в наборі даних, без необхідності його завантажувати й відкривати. Він має у стислій формі надавати інформацію про зміст набору даних.

Зокрема, в описі набору даних можна зазначити часовий період, яким він обмежується, згадати адміністративно-територіальні одиниці, яких він стосується, навести перелік найважливіших змінних, які в ньому вживаються.

Опис набору даних також може містити інформацію щодо методології збору та трансформації даних. Окрім того, в описі є сенс зазначити можливі способи застосування даних, наприклад, види аналізу, який дає змогу здійснювати набір даних.

В описі не варто наводити цитати з набору даних або ж нормативно-правових актів, пов’язаних із ним.

Приклад некоректного опису набору даних: Місто Київ – освітній, науковий, політичний, культурний, спортивний та економічний центр України, місто-лідер в Україні за кількістю народжених, часткою молоді в загальній чисельності населення, кількістю студентів і внутрішніх мігрантів, які обрали місто для життя і праці та є переважно молодими людьми. Усього в місті Києві станом на серпень 2015 року мешкало 2888,3 тис. осіб, з них близько 858 тис. осіб – молодь віком 14-35 років, що становить 30% від загальної чисельності населення міста. Зважаючи на високу частку молодого населення міста, високий відносно інших міст і регіонів держави рівень соціально-економічного розвитку, а також особливий статус Києва



та його суттєве значення для побудови процвітаючої України, у місті необхідно формувати і впроваджувати активну політику щодо інтелектуального, морального, фізичного розвитку молоді, реалізації її освітнього й творчого потенціалу, спортивних та інших досягнень.

Приклад коректного опису набору даних: Набір даних містить інформацію щодо Міської комплексної цільової програми «Молодь та спорт столиці» на 2016-2018 роки. Зокрема у наборі є паспорт програми, її бюджет за різними напрямками та підпрограмами, інформація про підпрограми підтримки молоді та розвитку спорту, основні напрями діяльності й перелік заходів у межах програми, а також результативні показники її виконання.

КЛЮЧОВІ СЛОВА

Ключові слова потрібні для того, щоб набори даних можна було об'єднати в тематичні групи, за якими їх легко знайти й ідентифікувати. Ключові слова також використовуються як елементи інтерфейсу Єдиного державного порталу відкритих даних: клік на певне ключове слово дає змогу отримати всі набори даних, в описі яких воно використовується.

Власне тому як ключові варто використовувати короткі слова, які в загальній формі описують зміст набору даних.

Не варто використовувати словосполучення чи фрази, вузькі і специфічні терміни. Не треба без потреби запроваджувати нові ключові слова для опису набору даних. Насамперед варто використовувати ключові слова вже наявні на Єдиному державному порталі відкритих даних.

Приклад некоректних ключових слів: наказ, запити, запити на отримання публічної інформації, Інструкція про порядок обчислення та внесення платежів за спеціальне використання рибних та інших водних живих ресурсів, Податкова знижка, сума витрат
Класифікація товару(ів), Основні правила інтерпретації УКТЗЕД, примітки до розділів, примітки до груп, код(и) товару(ів), товарна група, позиція, підпозиція, категорія, підкатегорія, Переміщення через митний кордон, Товари, Митна вартість, Контроль, Посадові особи,



Рекомендації, екологічний податок, побутові відходи, Право на спадщину та/або посвідчені договори дарування, методичні рекомендації щодо організації та проведення органами державної фіскальної служби зустрічних звірок, обміну податковою інформацією при здійсненні податкового контролю використання води юридичними особами, фізичними особами-підприємцями та платниками єдиного податку для задоволення виключно власних питних і санітарно-гігієнічних потреб Перелік відомостей, які містять службову інформацію Державний реєстр реєстраторів розрахункових операцій (Реєстр РРО).

Приклад коректних ключових слів: НПА, наказ, постанова, рішення, ДФС.

СТРУКТУРА НАБОРУ ДАНИХ

У Постанові Кабінету Міністрів України №835 (зі змінами) структура набору даних визначається як сукупність метаданих, що містить опис складу (елементів) набору даних, їхній формат, параметри та призначення. Якщо говорити простіше, структура набору даних – це перелік усіх змінних, що у ньому є, з розшифруванням значення та зазначенням типу даних для кожної змінної.

Структура набору даних, як і опис, потрібна для того, щоб користувачі мали можливість зрозуміти, якого роду інформація є в наборі, без потреби завантажувати його та відкривати. Водночас це свого роду словник набору даних, який дає змогу користувачам краще його інтерпретувати.

Припустимо, що ми маємо такий набір даних:

region	year	population
Київ	2015	2888.0
Київ	2016	2906.6
Київ	2017	2925.8



Структура набору даних у такому разі буде мати такий вигляд:

variable	description	type
region	регіон (область, місто)	string
year	рік, за який подається інформацій	integer
population	чисельність наявного населення станом на перше січня, тис. осіб	numeric

В описі структури набору даних також можна зазначати, які значення може набувати та чи інша змінна. У такому разі структуру можна використовувати для валідації набору даних, тобто перевірки його коректності.

Наприклад, ви можете зазначити, що значення певної змінної мають бути унікальними, а значення іншої – можуть бути порожніми. Або ж ви можете вказати, що довжина значення якоїсь із текстових змінних не може перевищувати певну кількість знаків.

Структуру набору даних варто публікувати в машиночитаних форматах CSV, JSON, XML чи RDF. Для публікації структури набору даних категорично не підходять DOC(X), PDF, PNG і аналогічні їм файлові формати.

Для створення структури набору даних ви можете скористатись онлайн-сервісом [Data Package Creator](#), створеним міжнародною організацією Open Knowledge Foundation. Він дає змогу генерувати структуру набору даних у машиночитаному форматі JSON, на основі файлу і даними, а також зазначити для кожної змінної тип даних і опис.

Якщо ви не завантажите структуру під час додавання набору даних у машиночитаному форматі (CSV, JSON, XML) на Єдиний державний портал відкритих даних, механізми порталу запропонують вам згенерувати її вручну. Вам буде запропонована проста веб-форма, у яку можна додати перелік усіх змінних, що є в наборі даних, а також їхній опис.



ООНОВЛЕННЯ НАБОРІВ ДАНИХ

Щодня центральні органи виконавчої влади й органи місцевого самоврядування продукують величезну кількість інформації. Дані в публічних реєстрах постійно змінюються та доповнюються. Отже, потрібно оновлювати й набори даних, оприлюднені на Єдиному державному порталі відкритих даних. Цей розділ має на меті надати роз'яснення щодо порядку оновлення наборів даних.

Розпорядники можуть визначати періодичність оновлення даних самостійно.

Згідно з Постановою Кабінету Міністрів України №835 (зі змінами), розпорядники інформації можуть самостійно визначати періодичність оновлення наборів даних на Єдиному державному порталі відкритих даних. Залежно від ситуації, вона може варіюватися від “відразу після внесення змін” – у разі оприлюднення даних за допомогою інтерфейсу прикладного програмування – до “щороку”.

Періодичність оновлення даних на Єдиному державному порталі відкритих даних може залежати від частоти оновлення самого набору даних у розпорядника, наявності технічних і кадрових ресурсів для забезпечення оновлення наборів даних, а також наявності конкретного запиту на оновлення даних від користувачів.

Єдиний державний портал відкритих даних має інструменти для оновлення наборів даних.

Під час завантаження нової версії файлу на Єдиний державний портал відкритих даних, стара версія автоматично буде видалена і замінена новою за умови, що обидва файли мають однакові ім'я та розширення.

Якщо ви не плануєте зберігати на Єдиному державному порталі всі попередні версії набору даних, тоді під час завантаження нової версії файлу:



- Не потрібно реєструвати новий набір даних, а достатньо просто додати нову версію файлу до вже наявного набору даних;
- Назва нового файлу має бути такою ж, як і назва старого;
- Стару версію файлу не потрібно видаляти, вона автоматично буде замінена новою.

Якщо ж ви маєте на меті надати користувачам кілька версій одного набору даних, тоді під час завантаження нової версії файлу:

- Не потрібно реєструвати новий набір даних, а достатньо просто додати нову версію файлу до вже наявного набору даних;
- Назва нового файлу має відрізнятися від назви старого;
- Стару версію файлу не потрібно видаляти.

Тож, якщо ви з певною періодичністю оновлюєте набір даних, який містить, наприклад, оперативні дані про баланс газу газосховищах, варто реєструвати один набір із загальною назвою.

Неправильно: Реєструвати окремий набір даних для кожної версії: “Баланс газу 2018-01-01”, “Баланс газу 2018-01-02” тощо.

Правильно: Реєструвати один набір даних із назвою “Баланс газу”.

Якщо ви не маєте на меті зберігати кілька версій одного файлу на Єдиному державному порталі відкритих даних, тоді:

Неправильно: Публікувати в наборі даних різні версії одного файлу з різними назвами. Наприклад, “balans_2018_01_01.csv”, “balans_2018_01_02.csv”, “balans_2018[1].csv”, “balans_2018[2].csv” і т. д.

Правильно: Публікувати всі версії одного файлу з однаковою назвою, наприклад, “balans_2018.csv”.



Якщо ж ви плануєте надавати користувачам доступ до різних версій файлу, тоді:

Неправильно: Публікувати всі версії одного файлу з однаковою назвою, наприклад, “balans_2018.csv”.

Правильно: Публікувати в наборі даних різні версії одного файлу з різними назвами відповідно до певної системи іменування версій. Наприклад, Наприклад, “balans_2018_01_01.csv”, “balans_2018_01_02.csv”.

ВИКОРИСТАННЯ НАБОРІВ ДАНИХ

Публікація інформації у форматі відкритих даних має на меті уможливити її швидке оброблення автоматизованими засобами. Маючи дані у структурованому та машиночитаному форматі, журналісти зможуть створити нові цікаві історії з картами та візуалізаціями, науковці – дослідити трансформації в політиці, економіці чи суспільстві, аналітики – розробити рекомендації щодо впровадження реформ і політик, активісти – контролювати виконання органами влади своїх обов’язків, бізнес – оптимізувати свою операційну діяльність, розробники – створити корисні сервіси та застосунки.

Інакше кажучи, відкриті дані існують передусім для використання. Без використання відкриття даних втрачає сенс.

Зрозуміло, що розпорядник даних не може нікого **змусити** використовувати свої дані. Однак він може **спонукати** та **заохотити**. Є декілька способів це зробити.



ПУБЛІКАЦІЯ ЯКІСНИХ НАБОРІВ ДАНИХ, ЯКІ Є ПРЕДМЕТОМ СУСПІЛЬНОГО ІНТЕРЕСУ

Ніщо так не стимулює використання даних, як наявність деталізованих, актуальних і якісних наборів даних, які є предметом високого суспільного інтересу та на які є запит у громадськості.

Традиційно **даними, які є предметом суспільного інтересу**, вважаються дані про бюджети всіх рівнів, публічні фінанси та державні закупівлі, законодавство, реєстр компаній і бенефіціарів, земельний кадастр і адміністративні кордони, якість повітря й води та прогноз погоди. Саме ці дані можуть стати основою аналітичного дослідження, журналістського розслідування чи сервісу.

Деталізованими наборами даних можна вважати такі, що містять неузагальнену, неагреговану інформацію. Наприклад, якщо йдеться про дані щодо якості повітря, то деталізованими можна вважати дані щоденних (або й навіть щогодинних) спостережень із кожної станції, а не дані про середній рівень забруднення повітря в місті за місяць.

Для користування також важливо, щоб дані були **актуальними** та **оновлюваними**. Навряд хтось захоче робити сервіс на основі даних за позаминулий рік. Так само навряд хтось наважиться робити сервіс на даних, у регулярності оновлення яких немає впевненості.

І нарешті дані мають бути **якісними**. Під цим можна мати на увазі як структурованість та машиночитаність даних, так і наявність словників і метаданих. Якщо користувачі не зможуть розібратись у опублікованому наборі даних через відсутність словників чи метаданих, він лежатиме невживаним на Порталі відкритих даних або буде хибно витлумаченим. Аби не допустити цього, потрібно публікувати якомога більше супровідної інформації про набір даних, яка допоможе користувачам краще зрозуміти та використовувати його.



КОМУНІКАЦІЯ ІЗ КОРИСТУВАЧАМИ

Продуктивна комунікація розпорядника інформації з користувачами (журналістами, аналітиками, розробниками, представниками бізнесу) також здатна заохотити їх до використання даних.

Комунікацію з користувачами можна умовно розділити на три етапи: до, під час та після оприлюднення даних.

До оприлюднення потрібно ідентифікувати попит користувачів на дані. Виявити, які саме набори даних та в якому вигляді мають найбільший потенціал для використання.

Щоб ідентифікувати попит на дані, можна провести онлайн-опитування, організувати зустріч або круглий стіл із зацікавленими особами чи організаціями, або ж влаштувати ідеатон – подію, під час якої потенційні користувачі та представники розпорядника даних обговорюють ідеї проектів на основі даних або проблеми, які можна вирішити за допомогою даних.

Це дає змогу зосередити зусилля розпорядника даних на найбільш цінних наборах, таких, що гарантовано будуть використані для журналістського розслідування, аналітичного дослідження чи соціально корисного сервісу. Розпорядник зможе підготувати ці набори даних до публікації у такому вигляді та форматі, у яких вони потрібні користувачам, а також визначити для них оптимальну частоту оновлення.

Під час оприлюднення варто донести інформацію про цінний набір даних до всіх зацікавлених осіб: поширити інформацію про нього на офіційній сторінці розпорядника, або/та у соціальних мережах, профільних спільнотах.

Значна частина наборів даних залишаються невикористаними з тієї простої причини, що потенційні користувачі не знають про їхнє існування. Розпорядник може і має докласти зусиль до поширення інформації про оприлюднений набір даних, щоб посприяти залученню журналістів, аналітиків та розробників до його використання.



Зрозуміло, що ця стратегія має застосовуватись лише для цікавих та якісних наборів даних. Якщо ви публікуєте десятки текстових документів на день, не варто повідомляти користувачам про кожен із них.

Після оприлюднення набору даних потрібно підтримувати зв'язок із користувачами, реагувати на їхні коментарі та побажання щодо опублікованих даних, виявляти проекти, розслідування чи дослідження на основі цих даних та публічно заохочувати їх авторів.

Деякі розпорядники публікують на своїх сайтах чи сторінках у соціальних мережах інформацію про проекти, зроблені на основі їхніх даних. Це слугує визнанням для авторів проектів і водночас популяризує дані та заохочує інших до їх дослідження.

ПРОВЕДЕННЯ ХАКАТОНІВ І КОНКУРСІВ

Для заохочення використання найбільш цінних і якісних наборів даних можна організувати хакатон чи конкурс проектів на основі відкритих даних.

Хакатон – це зазвичай дводенний захід, під час якого учасники – журналісти, аналітики, програмісти, дизайнери – працюють над вирішенням якогось завдання або розробляють проект на основі даних.

Хакатон дає змогу згуртувати навколо якоїсь проблеми (або набору даних) активних і зацікавлених людей, а також за короткий термін отримати прототип соціально корисного сервісу, аналітичного дашборду або чорновий варіант журналістського розслідування. У сфері відкритих даних хакатони є звичною справою. В Україні щороку проводять близько 10 хакатонів. На одному з них спеціалісти з data science розробляли алгоритми ідентифікації підозрілих закупівель у системі Prozorro, а на іншому активісти та розробники створювали антикорупційні інструменти для Міністерства інфраструктури.

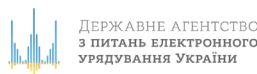


Хакатони добре працюють і на місцевому рівні – під час одного з хакатонів у Львові кілька груп програмістів працювали над розробкою сервісів для жителів міста.

Конкурси проектів на основі відкритих даних – не менш дієвий спосіб залучення користувачів. На відміну від хакатонів, конкурси можуть тривати значно довше ніж два дні, а також проводяться онлайн.

Як хороший приклад залучення користувачів до дослідження даних можна назвати конкурс Є-Рослідження, організований командою Єдиного порталу використання публічних коштів. Результатом цього конкурсу стала поява десятків розслідувань про використання бюджетних коштів, як на національному, так і на місцевому рівні.

Також варто відзначити масштабні конкурси проектів на основі відкритих даних [EGAP Challenge](#) та [Open Data Challenge](#). Результатом кожного з цих конкурсів стала поява корисних сервісів на основі відкритих даних, таких як Opendatabot або ж “Суд на долоні”. Ба більше, унаслідок проведення цих конкурсів сотні людей по всій країні долучилися до роботи з відкритими даними.



Матеріал підготовлено в рамках проекту «Прозорість та підзвітність в державному управлінні та послугах» / TAPAS. Створення цього матеріалу стало можливим завдяки підтримці американського народу, наданій через Агентство США з міжнародного розвитку (USAID) та за фінансової підтримки UK aid уряду Великої Британії. Зміст матеріалу не обов'язково відображає погляди Агентства USAID або уряду США чи офіційну політику уряду Великої Британії.