

ЗАТВЕРДЖЕНО

Наказ Державної служби
статистики

09 червня 2026 року № 148

**Порядок
використання публічної інформації у формі
відкритих даних щодо споживчих цін в Україні
для цілей офіційної статистики**

1. Вступ

1. Державна служба статистики України (Держстат) упроваджує сучасні методи збору, оброблення й аналізу цінової інформації для забезпечення повноти, своєчасності й достовірності статистичних спостережень за динамікою споживчих цін. Розвиток цифрової економіки, поширення електронної комерції та онлайн-платформ продажу товарів і послуг зумовлюють розширення використання джерел адміністративних даних для статистичних цілей. Одним із таких джерел є публічна інформація у формі відкритих даних (далі – публічна інформація), оприлюднена у відкритому доступі в мережі Інтернет власниками вебсайтів, торговельних майданчиків, постачальниками товарів і послуг, державними органами влади, органами місцевого самоврядування та іншими юридичними особами незалежно від організаційно-правової форми або фізичними особами (далі – вебресурси) з метою ознайомлення невизначеного кола користувачів з інформацією про ціни, тарифи й умови продажу товарів і послуг. Така інформація є доступною без авторизації чи спеціального дозволу, може бути технічно зібрана автоматизованими засобами без порушення умов користування вебресурсами або авторських прав і використовується виключно для статистичних цілей.

2. Під час збору публічної інформації не обробляються персональні дані чи будь-які відомості, що можуть ідентифікувати фізичних осіб та їх діяльність; збір обмежується лише тими відомостями, які власники ресурсів свідомо зробили відкритими для публічного поширення з метою інформування споживачів про свої товари, послуги або встановлені тарифи.

3. Метою цього Порядку є забезпечення прозорості, етичності та законності використання працівниками органів державної статистики публічної інформації у статистичних цілях, зокрема для розрахунку індексу споживчих цін, а також інформування користувачів офіційної державної статистичної інформації про те, що дані, публічно оприлюднені в Інтернеті, можуть бути зібрані та використані

Держстатом для цілей офіційної статистики, зокрема для розрахунку індексу споживчих цін. Це стосується всіх видів інформації про ціни, тарифи й умови продажу товарів і послуг, які власники вебресурсів свідомо зробили публічними.

4. Дія цього Порядку поширюється на департамент координації процесу збирання даних, департамент статистики цін апарату Держстату та державну установу «Центр обробки статистичних даних» під час проведення державних статистичних спостережень.

5. Держстат діє на засадах відкритості, добросовісності й поваги до прав усіх учасників інформаційного середовища. Публічна інформація збирається виключно в межах публічно оприлюднених відомостей, без втручання в роботу вебресурсів, без збору персональних даних і з дотриманням основних принципів офіційної статистики.

2. Підстави та правова база

1. Держстат збирає, обробляє та використовує публічно доступну інформацію у статистичних цілях відповідно до законодавства України та міжнародних статистичних стандартів та рекомендацій.

2. Правовою основою використання таких даних є:

1) Закон України «Про офіційну статистику», який визначає завдання органів державної статистики щодо збирання, узагальнення й оприлюднення статистичної інформації, необхідної для оцінки соціально-економічних процесів у державі;

2) закони України «Про інформацію», «Про захист інформації в інформаційно-комунікаційних системах», «Про електронні документи та електронний документообіг» та постанова Кабінету Міністрів України від 29 березня 2006 року № 373 «Про затвердження Мінімальних вимог до захисту інформаційних, електронних комунікаційних, інформаційно-комунікаційних та технологічних систем», що встановлюють загальні принципи створення, збирання, зберігання, поширення та використання інформації, зокрема з відкритих джерел;

3) Закон України «Про авторське право і суміжні права» та Цивільний кодекс України, які гарантують повагу до прав власників вебресурсів, структури та змісту баз даних.

3. Міжнародні статистичні стандарти та рекомендації офіційної статистики, зокрема Основні принципи офіційної статистики ООН (Fundamental Principles of Official Statistics), Кодекс практики європейської статистики (European Statistics Code of Practice), Регламент Європейського Парламенту та Ради (ЄС) № 223/2009

від 11 березня 2009 року про європейську статистику, стандарт SDMX для обміну даними та метаданими та рекомендації Євростату щодо використання альтернативних джерел даних (вебскрейпінг, сканерні дані, транзакційні бази) для цілей вимірювання інфляції.

4. Держстат збирає публічну інформацію виключно з тих вебресурсів, які оприлюднюють дані про товари, послуги та тарифи з метою публічного інформування споживачів, і лише в обсягах, необхідних для забезпечення повноти охоплення явища та якості офіційної державної статистичної інформації.

3. Науково-технічні принципи та професійна етика збору та використання публічно доступної інформації

1. Держстат дотримується стандартів етичності, добросовісності та професійної відповідальності у процесі збору й використання публічно доступної інформації. Застосування автоматизованих технологій збору (вебскрейпінгу) здійснюється виключно в межах законних, технічно безпечних та соціально відповідальних практик, узгоджених із міжнародними рекомендаціями Євростату, ОЕСР і ООН.

2. Основними принципами збору та використання публічно доступної інформації є:

1) законність і добросовісність, а саме збирання інформації лише з відкритих джерел, без авторизації, платних підписок або технічних обмежень, у повній відповідності до законодавства України. Держстат поважає права власників вебресурсів, дотримується умов користування сайтами, зокрема правил robots.txt, ліцензійних обмежень, і не завдає шкоди працездатності, доступності чи репутації вебресурсів;

2) технічна відповідальність та ідентифікація, що включає всі автоматизовані запити Держстату, має унікальний ідентифікатор (User-Agent), що дозволяє адміністраторам сайтів легко визначати статистичну активність у лог-файлах серверів. Приклад ідентифікатора: Ukrstat Statistics Bot v1.0.;

3) мінімізація втручання та захист приватності передбачає, що збір обмежується виключно необхідною інформацією про ціни, тарифи чи умови продажу, без зайвих технічних або контекстних відомостей. Із процесу збору повністю виключено будь-які метадані, які можуть містити персональні дані фізичних осіб, контактні або ідентифікаційні відомості;

4) пропорційність і періодичність збору передбачають, що запити до вебресурсів виконуються з урахуванням технічного навантаження та принципу мінімального втручання. Інформація збирається не частіше, ніж це необхідно для

фіксації реальних змін у цінах і тарифах, тобто з мінімально можливою періодичністю, що гарантує актуальність даних. За можливості перевага надається відкритим інтерфейсом програмування застосунків (прикладний програмний інтерфейс) (англ. application programming interface) (далі – API), прайс-листам та іншим публічним джерелам, які підтримують машинний доступ і не створюють додаткового навантаження на ресурси;

5) поєднання методів збору передбачає, що Держстат застосовує як вебскрейпінг, так і прямий обмін даними із власниками вебресурсів. Для великих сайтів або в разі регулярного оновлення інформації перевага надається організованим каналам обміну – відкритим API, файловому обміну чи іншим узгодженим технічним рішенням. Обидва методи розглядаються як взаємодоповнюючі, із пріоритетом ефективності, технічної безпеки та економічної доцільності;

б) прозорість і контроль якості означають, що всі етапи збору, очищення й оброблення публічної інформації документуються у внутрішніх процедурах Держстату для забезпечення відтворюваності процесів. Оскільки статистичне спостереження, яке використовує публічно доступні дані, на цей час має експериментальний характер, методологія перебуває на етапі вдосконалення. Публікуються лише узагальнені принципи, без розкриття технічних деталей або конкретних джерел, щоб не порушити стабільність і нейтральність процесу;

7) використання виключно у статистичних цілях передбачає, що зібрана публічна інформація використовується лише в межах офіційної статистики: для розрахунку індексу споживчих цін, удосконалення методів збирання інформації для проведення спостереження та розвитку статистичної інфраструктури;

8) контроль якості та протидія маніпуляціям, а саме для запобігання спотворення даних, застосовується багаторівнева система валідації. Інформація з одного джерела завжди зіставляється з даними з кількох інших, що дозволяє виявляти нетипові значення в ряду розподілу даних і сприяє достовірності результатів;

9) відкритість до діалогу та співпраці означає, що Держстат прагне підтримувати відкриту комунікацію із власниками вебресурсів. У разі запитань, технічних зауважень або пропозицій щодо ефективніших способів передачі даних (наприклад, через API), пропонуємо звертатися на адресу: datacollection@stat.gov.ua.

4. Методологічні засади збору публічно доступної інформації

1. Збір публічно доступної інформації здійснюється за стандартизованими процедурами, що забезпечують достовірність, репрезентативність і

порівнюваність статистичних результатів. Використання цифрових джерел даних – вебресурсів – дозволяє підвищити оперативність і забезпечити охоплення явища для спостережень за цінами й тарифами.

2. Процес збору передбачає:

1) визначення джерел отримання інформації здійснюється шляхом ідентифікації вебресурсів, котрі регулярно публікують цінову або тарифну інформацію в публічному доступі. Моніторинг цін відбувається за сформованим переліком товарів і послуг відповідно до структури споживчого набору та Класифікації індивідуального споживання за цілями (КІСЦ) затвердженого наказом Держкомстату від 29 грудня 2007 року № 480;

2) відбір вебресурсів здійснюється на основі оцінки, надійності, стабільності та повноти інформації. Для кожного джерела фіксується тип, формат і частота оновлення цінових даних;

3) використання комбінованих методів збору передбачає, що публічна інформація отримується як шляхом вебскрейпінгу (автоматизованого зчитування відкритих сторінок), так і через прямі канали передачі даних – API, файловий обмін або шляхом укладання угод щодо взаємообміну інформаційними ресурсами. Вибір методу визначається технічними можливостями, економічною доцільністю та регулярністю оновлення даних;

4) технічна реалізація збору передбачає, що алгоритми вебскрейпінгу налаштовуються так, щоб мінімізувати кількість запитів до вебресурсів, дотримуватися правил robots.txt, не завдавати шкоди працездатності сайтів і збирати лише необхідні для офіційної статистики дані та метадані;

5) валідація й очищення передбачають, що зібрана інформація проходить автоматизовану перевірку на повноту, коректність і узгодженість. Виключаються дублікати, некоректні значення та спотворення, що виникають через зміни структури сторінок;

6) зіставлення та класифікація, а саме ціни, отримані з різних джерел, узгоджуються згідно з КІСЦ, а також з переліками товарів (послуг) - представників, визначеними відповідно до методології проведення статистичного спостереження;

7) інтеграція в систему статистичного спостереження передбачає, що оброблена інформація завантажується у внутрішні бази даних Держстату, де далі використовується для формування індивідуальних та зведених індексів.

5. Оброблення, перевірка та використання публічно доступної інформації

1. Публічна інформація, зібрана з відкритих джерел, проходить комплексну систему технічної та аналітичної перевірки, яка забезпечує її достовірність, узгодженість і придатність для статистичних розрахунків. Метою цього процесу є перетворення неструктурованих онлайн-відомостей про ціни й тарифи в упорядковану та знеособлену статистичну інформацію.

2. Етапи оброблення та контролю якості:

1) первинна перевірка, під час якої здійснюється автоматизоване тестування коректності форматів, відсутності помилок завантаження, дублікатів і нетипових значень;

2) фільтрація й очищення, де з отриманих даних вилючаються записи, що не містять цінкових даних, мають технічні або рекламні ознаки чи не відповідають визначеній меті спостереження;

3) контроль узгодженості передбачає, що ціни, отримані з різних джерел, перевіряються на логічну й аналітичну відповідність. У разі виявлення відхилень проводиться додаткова перевірка або повторний збір;

4) агрегація та знеособлення передбачають, що після валідації дані об'єднуються у зведену базу, у якій відсутня будь-яка інформація, що може ідентифікувати конкретне джерело;

5) внутрішній моніторинг (документування отриманої інформації), передбачає технічний і методичний контроль на кожному етапі, ведення журналу дій та історії змін для забезпечення повної відтворюваності процесів;

6) захист інформації передбачає, що Держстат забезпечує належний рівень захисту зібраної публічної інформації з дотриманням вимог законодавства України про офіційну статистику, у сферах захисту інформації, електронних довірчих послуг та електронного документообігу, рівень захисту якої повинен відповідати ступеню обмеження доступу до неї.

3. На всіх етапах оброблення – від збору до аналітичного узагальнення – застосовуються технічні та організаційні заходи, спрямовані на запобігання несанкціонованому доступу, копіюванню або використанню даних. До таких заходів належать обмеження доступу до первинної інформації, автентифікація, контроль прав доступу, а також використання засобів шифрування.

4. Працівники, залучені до процесів збору, оброблення й узагальнення інформації, зобов'язані дотримуватись принципу статистичної конфіденційності відповідно до Закону України «Про офіційну статистику».

5. Оновлення та повторна перевірка передбачає, що алгоритми збору та перевірки регулярно переглядаються, щоб ураховувати зміни у структурі, візуалізації та технологіях вебресурсів.

6. Публічна інформація використовується з урахуванням принципів офіційної статистики, визначених пунктом 3 частини першої статті 4 Закону України «Про офіційну статистику». Отримані дані застосовуються для статистичних цілей, зокрема для розрахунків індексу споживчих цін і вдосконалення статистичних методів розрахунку індексів споживчих цін.

7. Держстат не розкриває алгоритми, технічні рішення чи програмні засоби, що застосовуються під час збору й оброблення даних, оскільки така інформація не має практичного значення для користувачів, але може створити ризики порушення стабільності збору та ідентифікації джерел. Водночас усі зібрані відомості проходять перевірку, очищення й узагальнення, під час яких видаляються будь-які посилання на окремі джерела інформації.

Отже, вихідні статистичні показники не містять даних та метаданих щодо кожного учасника ринку та використовуються виключно у знеособленому вигляді.

6. Заключні положення

1. Держстат не втручається у зміст або структуру вебресурсів, з яких збирається публічна інформація, не впливає на політику ціноутворення чи діяльність суб'єктів господарювання, не дає оцінок окремим компаніям або ринкам.

2. Цей Порядок висвітлює засади збору та використання публічної інформації для цілей офіційної статистики, та покликаний підкреслити прозорість дій Держстату щодо використання публічної інформації у статистичних цілях, зміцнити довіру користувачів офіційної державної статистичної інформації до офіційної статистики.

Директор департаменту координації
процесу збирання даних апарату Держстату

Ольга ПРАВОСУД