



Міністерство освіти і науки України  
Київський національний університет будівництва і  
архітектури



Кафедра геоінформатики і фотограмметрії

## Оцінювання якості відкритих геопросторових даних

Виконала: студентка 4 курсу, групи ГІСТ-21  
Спеціальності: 193 «Геодезія та землеустрій»,  
Спеціалізації: «Геоінформаційні системи та технології»  
Вероніка БОЖКО  
Керівник: доц., PhD  
Данило КІНЬ  
Консультант: док. техн. наук, професор  
Юрій КАРПІНСЬКИЙ

Київ – 2025

## Актуальність дослідження



У сучасних умовах цифрової трансформації геопросторові дані стають критично важливими для містобудування, транспорту, екології та безпеки. Відкриті джерела, як OpenStreetMap, Google Maps, Esri, Microsoft Building Footprints, Bing Maps набувають популярності завдяки доступності й охопленню, але відповідність вимогам до точності, повноти та детальності є відкритим питанням.

Відсутність контролю за якістю таких даних створює ризики для просторового аналізу та проектування. Це зумовлює потребу у методиці оцінювання геопросторових даних, заснованій на порівнянні з еталонними моделями. Відповідно до Закону України «Про національну інфраструктуру геопросторових даних» та положення стандартів ISO, така перевірка є обов'язковою.



# МЕТА, ОБ'ЄКТ, ПРЕДМЕТ ДОСЛІДЖЕННЯ

- **Мета дослідження:**

Виконання оцінки якості відкритих геопросторових даних із використанням сучасних геоінформаційних технологій. Досягнення цієї мети сприятиме ефективнішому застосуванню геопросторових даних завдяки забезпеченню їхньої надійності та відповідності потребам користувачів для прийняття обґрунтованих рішень у різних сферах діяльності.

- **Об'єкт дослідження:**

Набори відкритих геопросторових даних.

- **Предмет дослідження:**

Якість наборів відкритих геопросторових даних.

## **Основні завдання:**

Проаналізувати нормативно-правове забезпечення щодо якості геопросторових даних та вимог до неї;

Дослідити міжнародний та вітчизняний досвід щодо виконання оцінки якості геопросторових даних;

Сформувати метадані вихідних наборів геопросторових даних проєкту та опрацювати їх;

Описати методику оцінки якості геопросторових даних;

Розробити концептуальну модель бази геопросторових даних для оцінювання якості даних;

Розробити логічну модель зазначеної бази геопросторових даних;

Реалізувати базу геопросторових даних проєкту у середовищі ArcGIS;

Виконати оцінку якості наборів відкритих геопросторових даних у середовищі ArcGIS;

# Нормативно-правове забезпечення щодо якості геопросторових даних



## Законодавство України

- Закон України «Про національну інфраструктуру геопросторових даних» (2020).
- Технічні вимоги та методи забезпечення інтероперабельності і сумісності наборів геопросторових даних та геоінформаційних сервісів (2021).
- Технічні вимоги до метаданих для наборів геопросторових даних і геоінформаційних сервісів національної інфраструктури геопросторових даних (2021).
- Про затвердження Порядку функціонування національної інфраструктури геопросторових даних (2021)



## Міжнародні та національні стандарти

- Директива 2007/2/ЄС про створення Інфраструктури просторової інформації у Європейському Співтоваристві (INSPIRE).
- ISO 19157:2013 «Географічна інформація – Якість даних».
- Open Geospatial Consortium (OGC) Standards.



## Практична реалізація

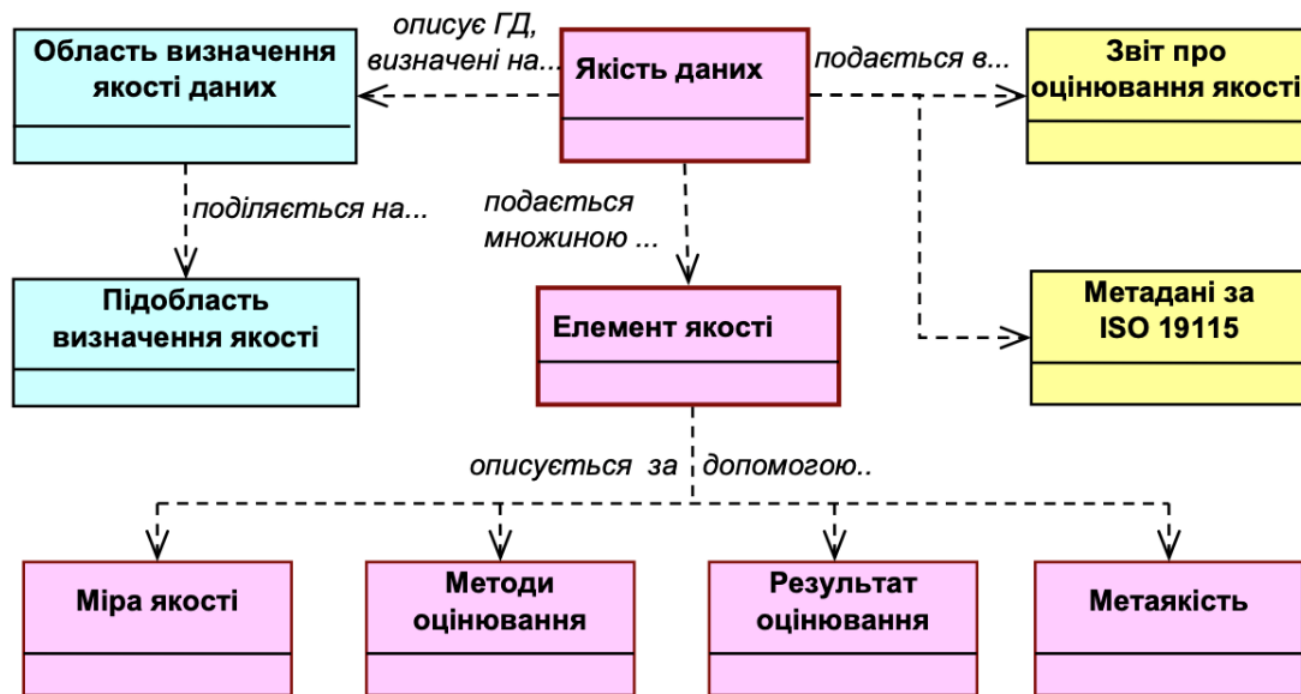
- Національний геопортал НІГД
- ГІС містобудівного кадастру на державному рівні
- ГІС Державного картографо-геодезичного фонду України
- Національна кадастрова система

# Ключові вимоги до якості відкритих геопросторових даних

У контексті оцінювання якості відкритих геопросторових даних ключовими характеристиками

виступають **повнота**, **детальність** та **позиційна точність**. Відповідно до концептуальної моделі якості геоданих (за ISO 19157), кожна з цих характеристик має своє місце у структурі якості: або як **елемент якості**, або як **міра оцінювання**.

- **Повнота** – **елемент якості**, описує наявність чи відсутність об'єктів, їх атрибутів чи відношень [1].
- **Детальність** – **елемент якості**, що характеризує геометричну складність об'єкта: кількість вершин, довжина сегментів, точність обрису.
- **Позиційна точність** – **елемент якості**, визначається як точність визначення координат місцеположення об'єктів у обраній референційній системі координат [1].



Концептуальна модель якості геоданих (ISO 19157, адаптовано проф. Лященком А.А.)

Модель відображає логіку оцінювання якості геопросторових даних згідно з ISO 19157. Систематизовано зв'язки між елементами якості (повнота, точність тощо), мірами, методами оцінювання та результатами, які фіксуються в метаданих та звітах.

[1] Основи створення інтероперабельних геопросторових даних / Ю.О. Карпінський та ін. – Київ: КНУБА, 2023. – 302 с.

# Аналіз останніх досліджень щодо оцінки якості геопросторових даних

## 1. Jingfeng Xia – «Metrics to Measure Open Geospatial Data Quality» (2012)

**Що оцінювалося:** Запропоновано класифікацію метричних характеристик якості відкритих геоданих – точність, повнота, узгодженість, актуальність тощо.

**Методи:** Комбіновані кількісні та якісні показники, створено концептуальну модель.

**Висновок:** Геопросторові дані мають унікальні властивості (координати, метадані), тому потребують спеціалізованих метрик оцінки.

**Враховано у роботі:** Сформовано перелік критеріїв якості (повнота, точність), орієнтація на ISO-стандарти.

## 2. Nayati Taştan , M.Orhan Altan– «Spatial Data Quality» (1999)

**Що оцінювалося:** Просторові показники (позиційні, тематичні), джерела похибок, моделі якості.

**Методи:** Розподіл похибок за джерелами (збір, обробка, зберігання тощо), введено поняття допустимого рівня якості (AQL).

**Висновок:** Якість геопросторових даних – це не лише точність, а й повнота, логічна послідовність, узгодженість, походження.

**Враховано у роботі:** Враховано одні із основних особливостей оцінки якості даних – відповідність, використання, суб'єктивність сприйняття.

## 3. Goma M. Dawod , Ibrahim E. Ascoura– «The Validity of Open-Source Elevations for Different Topographic Map Scales and Geomatics Applications» (2021)

**Що оцінювалося:** Якість відкритих цифрових моделей рельєфу (GDEM) у двох регіонах (Єгипет, Саудівська Аравія).

**Методи:** Порівняння висот з GDEM з опорними точками GCP, розрахунок RMSE (СКП).

**Висновок:** У рівнинних регіонах GDEM не можна використовувати в масштабах 1:25 000 і крупніше, у горбистих — не точніше ніж 1:50 000.

**Враховано у роботі:** Підхід до визначення придатності джерела за масштабом, відповідність точності масштабу плану/карти.

## 4. Milad Moradi, Stéphane Roche and Mir Abolfazl Mostafavi «Evaluating OSM Building Footprint Data Quality in Quebec» (2023)

**Що оцінювалося:** Повнота, точність положення та форми будівель OSM.

**Методи:** Порівняння з еталонними даними, аналіз кількості, площі, точності геометрії.

**Висновок:** Повнота зростає в залежності від міста (від 5% до 99%); позиційна точність в межах 2–3 м.

**Враховано у роботі:** Методика аналізу повноти та точності за площею і положенням.

# Опис методики оцінки якості геопросторових даних

---

Вибір та обґрунтування обраних елементів і мір якості відкритих геопросторових даних

---

Опис математичних моделей елементів та мір якості

---

Формулювання переліку наборів даних, якість яких оцінюється, та їх метаданих

---

Формулювання еталонних моделей для оцінки якості

---

Розроблення бази геопросторових даних для оцінки якості зазначених наборів

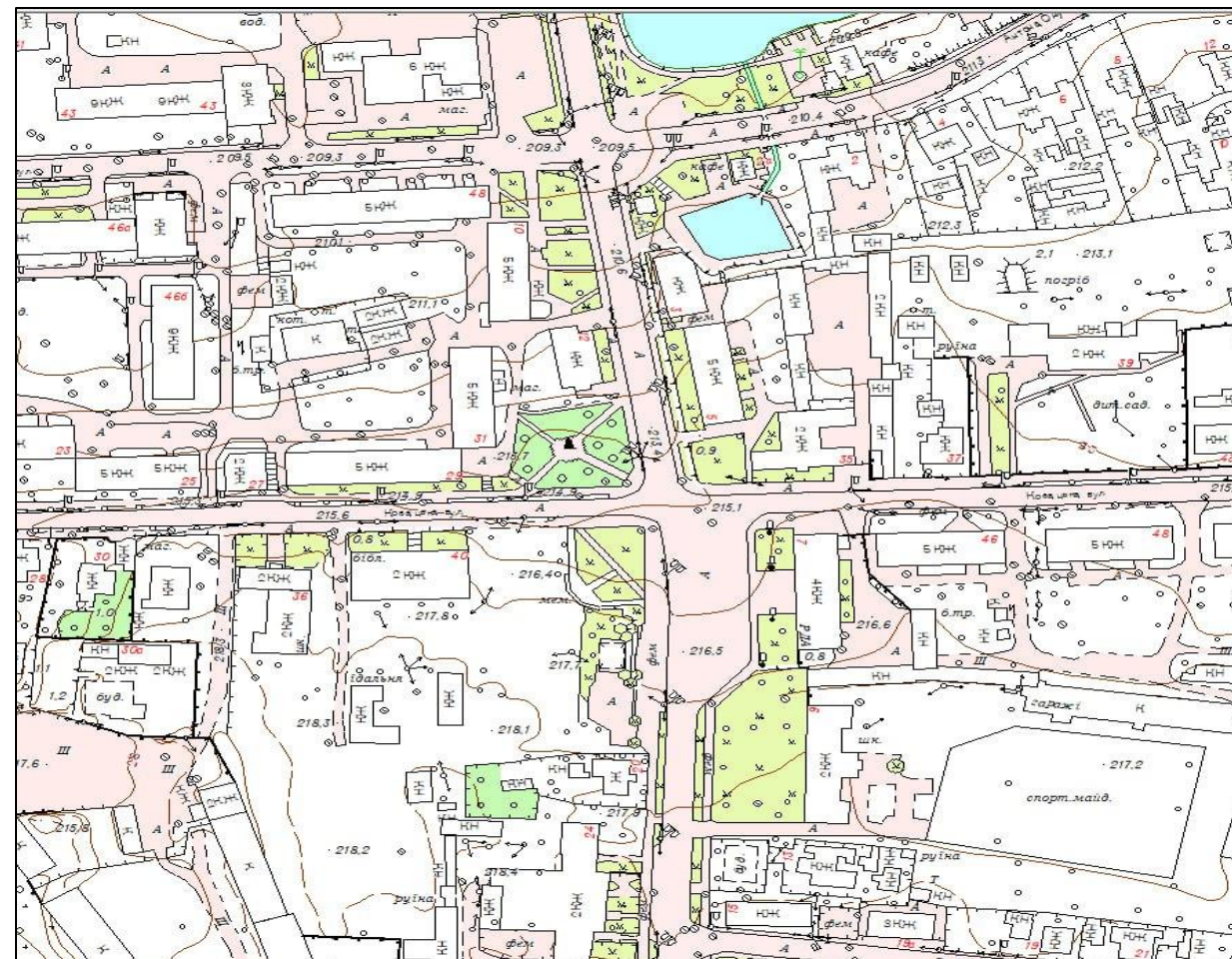
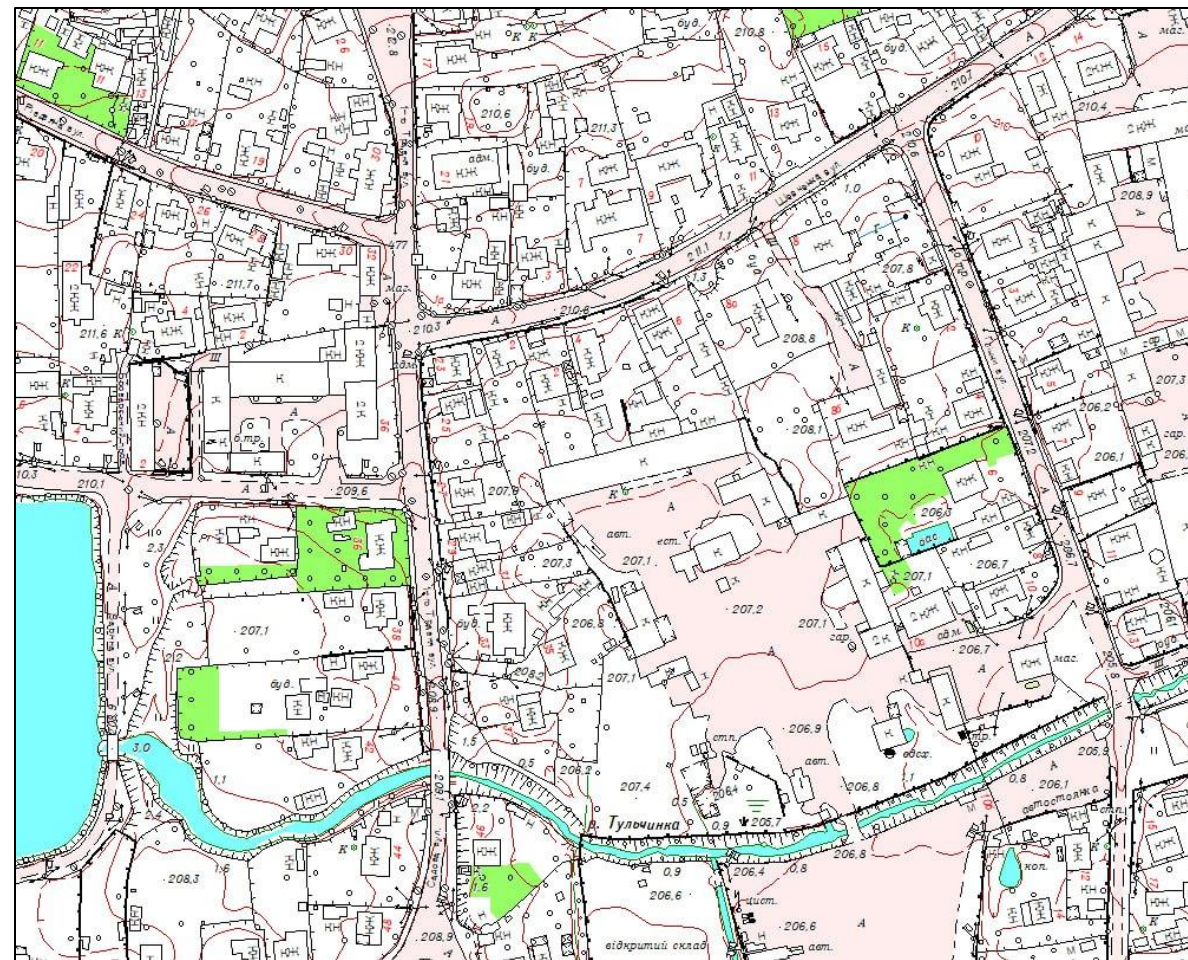
---

Виконання оцінки у геоінформаційному середовищі

---

Аналіз отриманих результаті , формування висновків та рекомендацій, щодо зазначених наборів

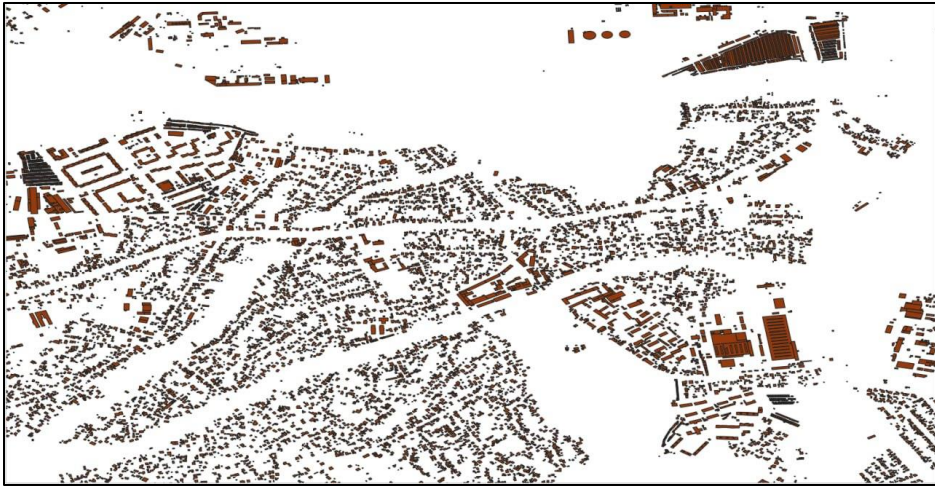
## Еталонні об'єкти на топографічних планах 1:2 000



**Тулчин** – місто, адміністративний центр Тульчинського району Вінницької області та Тульчинської міської громади

**Славута** – місто в Шепетівському районі Хмельницької області

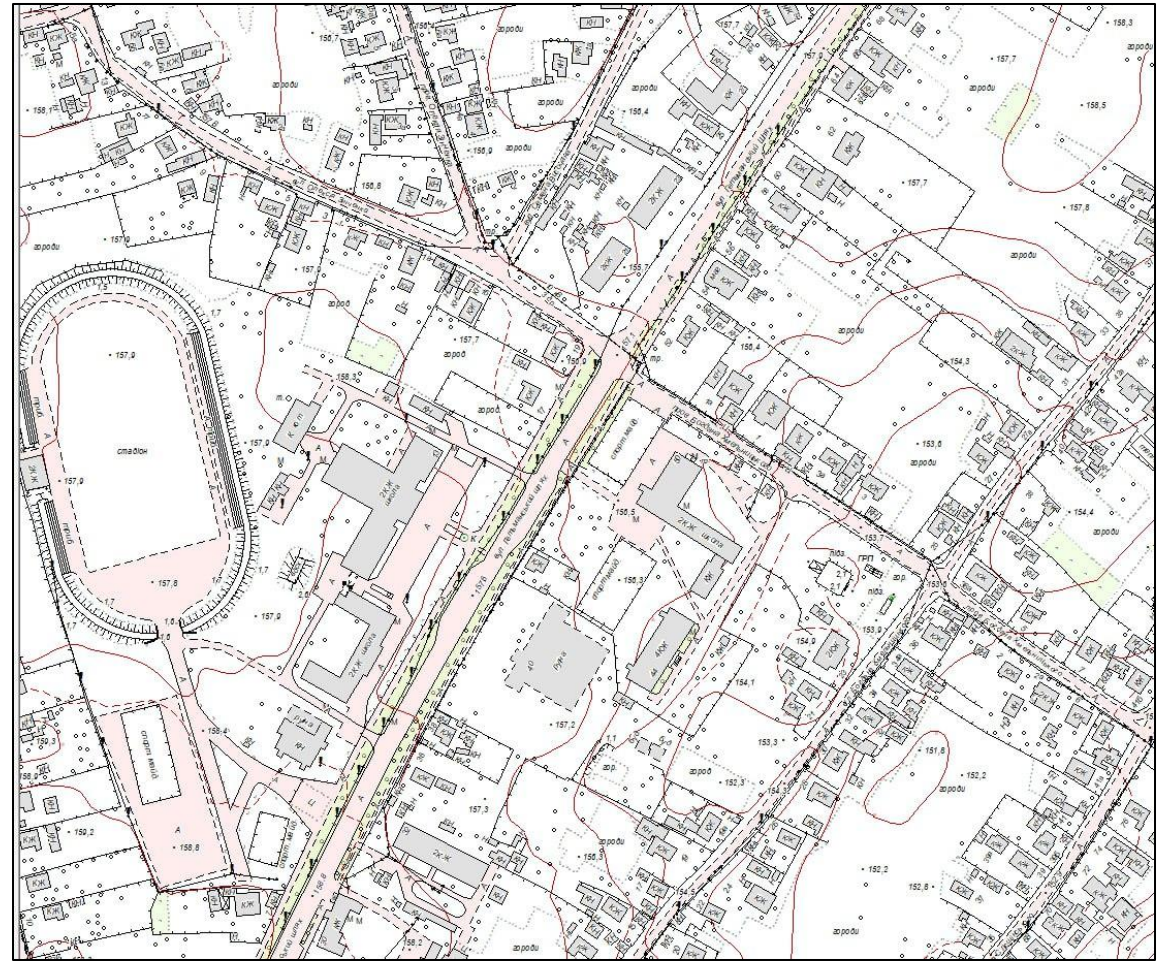
# Еталонні об'єкти на топографічних планах 1:2 000



**Пісочин** – селище Харківського району Харківської області



**Рогань** – адміністративний центр Роганської селищної громади Харківського району Харківської області



**Лисянка** – селище міського типу Черкаської області

## Набір відкритих геопросторових даних, що оцінюється



**OpenStreetMap (OSM)** – векторна база даних, створена спільнотою користувачів. Інформація про будівлі експортується за допомогою плагіна QuickOSM у форматі .osm і конвертується у .shp.

**Microsoft Bing Building Footprints** – набір векторних полігонів будівель, згенерованих автоматично з використанням нейромереж. Дані завантажуються з відкритого репозиторію GitHub у форматі GeoJSON.

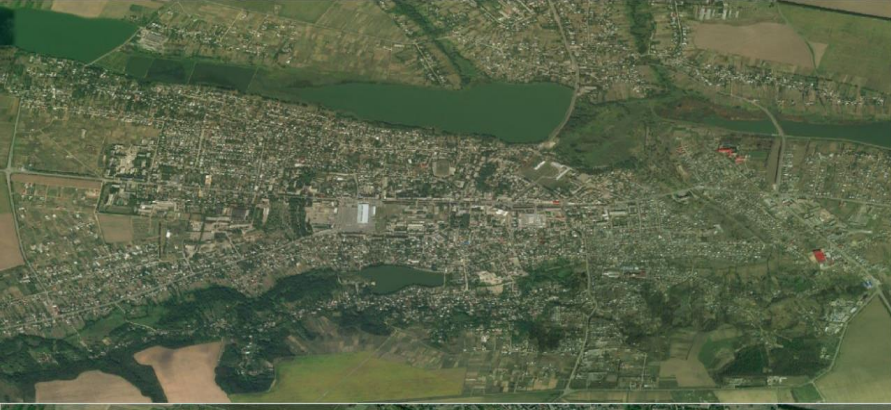
*Приклад на територію міста Славута.*

## Набір відкритих геопросторових даних, що оцінюється



**Google Satellite** – супутникова мозаїка високої роздільної здатності, інтегрована в QGIS через XYZ Tile Layer із використанням шаблону:

<http://mt0.google.com/vt/lyrs=s&hl=en&x={x}&y={y}&z={z}> .



**Bing Maps** – супутникові знімки Microsoft, підключені як растрова основа через шаблон XYZ:

<http://ecn.t3.tiles.virtualearth.net/tiles/a{q}.jpeg?g=1>



**Esri Imagery** – супутниковий сервіс від компанії Esri, підключається у вигляді XYZ Tile Layer через URL-шаблон:  
[https://server.arcgisonline.com/arcgis/rest/services/World\\_Imagery/MapServer/tile/{z}/{y}/{x}](https://server.arcgisonline.com/arcgis/rest/services/World_Imagery/MapServer/tile/{z}/{y}/{x}).

*Приклад на територію міста Славути.*

# Метадані набору відкритих геопросторових даних

№ з/п	Назва елемента метаданих та його позначка за ISO 19115-1:2014	Коментар	Значення метаданих
1	2	3	4
1	Унікальний ідентифікатор метаданих: (MD_Metadata.metadataIdentifier: MD_Identifier)	Унікальний ідентифікатор для метаданих, що формується програмними засобами підготовки метаданих	–
2	Назва ресурсу: (MD_Metadata.identificationInfo > MD_DataIdentification.citation > CI_Citation.title)	Характерна унікальна назва, що ідентифікує ресурс (набір даних або комплект наборів даних)	OpenStreetMap
3	Дата створення/оновлення ресурсу: (MD_Metadata.identificationInfo > MD_DataIdentification.citation > CI_Citation.date та CI_Citation.dateType)	за класифікатором CI_DateTypeCode	
		створення	2004-08-09T09:00:00
		оновлення	2025-02-09T21:21:00
4	Контактна інформація відповідальних за ресурс: (MD_Metadata.identificationInfo > MD_DataIdentification.PointOfContact > CI_Responsibility)	за класифікатором CI_RoleCode	
		постачальник	OpenStreetMap Foundation
		держатель	OpenStreetMap Foundation
5	Просторове охоплення: (MD_Metadata.identificationInfo > MD_DataIdentification.extent > EX_Extent.geographicElement > EX_GeographicExtent > EX_GeographicBoundingBox або EX_BoundingPolygon або EX_GeographicDescription)	Опис просторового охоплення території даними ресурсу	EPSG:3857 - 20037508.3427892439067364,- 20037508.3427892476320267 : 20037508.3427892439067364,200375 08.3427892476320267
6	Мова ресурсу: (MD_Metadata > MD_DataIdentification.defaultLocale > PT_Locale)	Код мови та код набору символів, що використовуються в ресурсі. За замовчуванням: код мови: Ukr; код набору символів: utf-8.	UTF-8

На слайді подано приклад заповнення метаданих для набору відкритих геопросторових даних **OpenStreetMap** відповідно до стандарту **ISO 19115-1:2014**.

Це таблиця містить основну інформацію про ресурс (назву, дату створення, відповідальних осіб) і використовується для документування джерела даних під час оцінювання їх якості.

# Метадані набору відкритих геопросторових даних

7	Тематична категорія ресурсу: (MD_Metadata.identificationInfo > MD_DataIdentification.topicCategory > MD_TopicCategoryCode)	за класифікатором базових категорій, що визначені в ISO 19115-1:2014, який може доповнюватися.	Межі; Рельєф; Базові зображення поверхні Землі;
8	Просторове розрізнення: (MD_Metadata.identificationInfo > MD_Identification.spatialResolution > MD_Resolution.equivalentScale, MD_Resolution.distance, MD_Resolution.vertical або MD_Resolution.angularDistance або MD_Resolution.levelOfDetail)	Числовий масштаб (рівень деталізації)	Залежить від джерел даних, території
9	Тип ресурсу: (MD_Metadata.metadataScope > MD_Scope.resourceScope)	за класифікатором MD_ScopeCode	dataset
10	Стислий опис ресурсу: (MD_Metadata.identificationInfo > MD_Identification.abstract)	Реферативний опис, призначення, структури, складу та змісту набору геопросторових даних та інші відомості про ресурс.	OpenStreetMap — це глобальна вільна картографічна платформа, створена для збору, збереження та поширення геопросторових даних, доступних для всіх. Вона призначена для використання в навігаційних системах, геоінформаційних системах (ГІС), містобудівній діяльності, наукових дослідженнях, веб-картах і мобільних додатках. Склад та зміст набору: адміністративні межі (країни, області, громади); транспортна інфраструктура (вулиці, залізниці, мости, тунелі); гідрографія (річки, озера) житлова та комерційна забудова зелені зони та землекористування інфраструктурні об'єкти (школи, лікарні, АЗС, магазини) мобільна інфраструктура (велодоріжки, зупинки, маршрути); історичні, культурні та туристичні об'єкти
11	Походження ресурсу: (MD_Metadata > resourceLineage > LI_Lineage)	Опис джерел ресурсу та основних технологічних процесів, що використовувалися для виробництва ресурсу.	Багато даних в OSM вручну векторизовані з аерофотозйомки чи супутникових зображень високої роздільної здатності (наприклад, Bing, Maxar). Часто прив'язані до GPS-треків, що забезпечує високу точність. У містах рівень деталізації іноді перевищує 1 м (особливо, якщо активно картографують місцеві волонтери).

# Математичні моделі елементів та мір якості геопросторових даних

## 1. Позиційна точність

Розраховується за допомогою формули середньоквадратичної похибки:

Формула Гаусса (коли істинне значення відоме):

$$m = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \Delta_i^2}$$

де  $i = 1 \dots n$ ;

$n$  – кількість вимірів;

$\Delta_i$  – похибка  $i$ -го вимірювання, яка визначається як різниця між виміряним значенням координати та еталонним (істинним) значенням координати.

$$\Delta_i = x_i^{\text{оцін}} - x_i^{\text{еталон}}$$

Формула Бесселя (коли істинне значення невідоме):

$$m = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n \delta_i^2}$$

де  $i = 1 \dots n$ ;

$\delta_i$  – різниця між виміряним значенням та середнім арифметичним значенням даної величини.

$$\delta_i = x_i - \bar{x}$$

де  $x_i$  – виміряне значення координати;

$\bar{x}$  – середнє арифметичне всіх виміряних значень.

# Математичні моделі елементів та мір якості геопросторових даних

## 2. Повнота

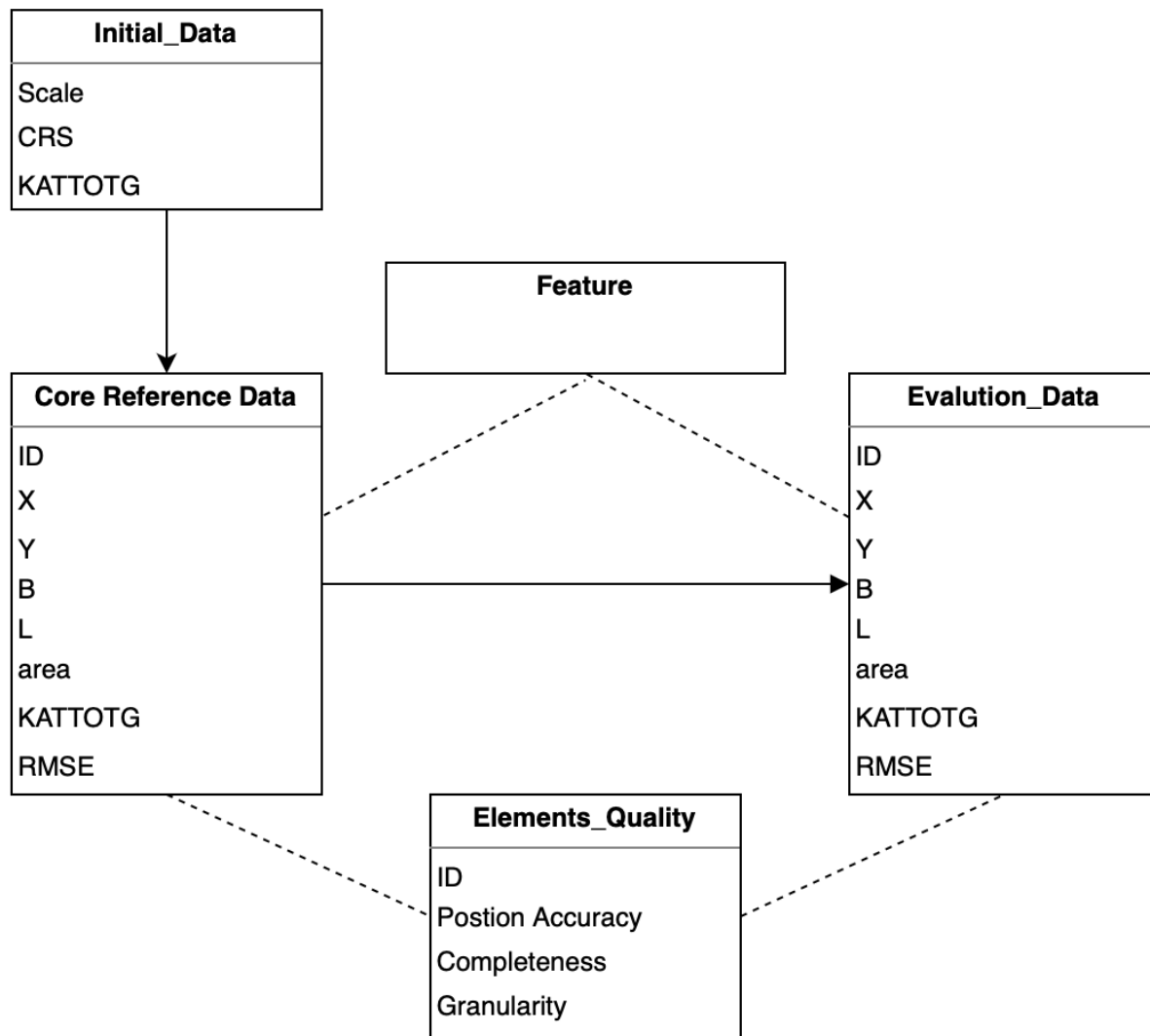
Для оцінювання повноти використано мінімальні розміри полігональних об'єктів згідно з масштабом. Будівлі, площа яких потрапляє у відповідний інтервал, вважаються такими, що відповідають вимогам до масштабу.

Масштаб	Інтервал допустимих площ будівель (м <sup>2</sup> )
1:500	[0,25 ; 0,99)
1:1000	[1,00 ; 3,99)
<b>1:2000</b>	<b>[4,00 ; 24,99)</b>
1:5000	[25,00 ; 99,99)
1:10000	[100,00 ; 624,99)
1:25000	[625,00 ; 2499,99)
1:50000	[2500,00 ; 9999,99)

## 3. Детальність

Максимальна відстань між точками векторних моделей контурів об'єктів залежно від масштабу топографічної зйомки, що визначає роздільну здатність векторних моделей контурів об'єктів, **не повинна перевищувати: у масштабі 1:2000 – 40 м.**

Масштаб	Максимальна відстань між точками векторних моделей контурів об'єктів
1:500	15 м
1:1000	20 м
<b>1:2000</b>	<b>40 м</b>
1:5000	60 м



## Концептуальна модель бази геопросторових даних

На слайді представлено концептуальну модель, яка відображає взаємозв'язки між еталонними (Core Reference Data) та оцінюваними (Evaluation Data) геоданими.

Модель дозволяє фіксувати координати, площі, коди територій (КАТОТТГ) і позиційні похибки (RMSE). Через таблицю Elements\_Quality здійснюється оцінка за критеріями: позиційна точність, повнота та детальність.

Така структура забезпечує системний підхід до оцінювання якості відкритих геопросторових даних.

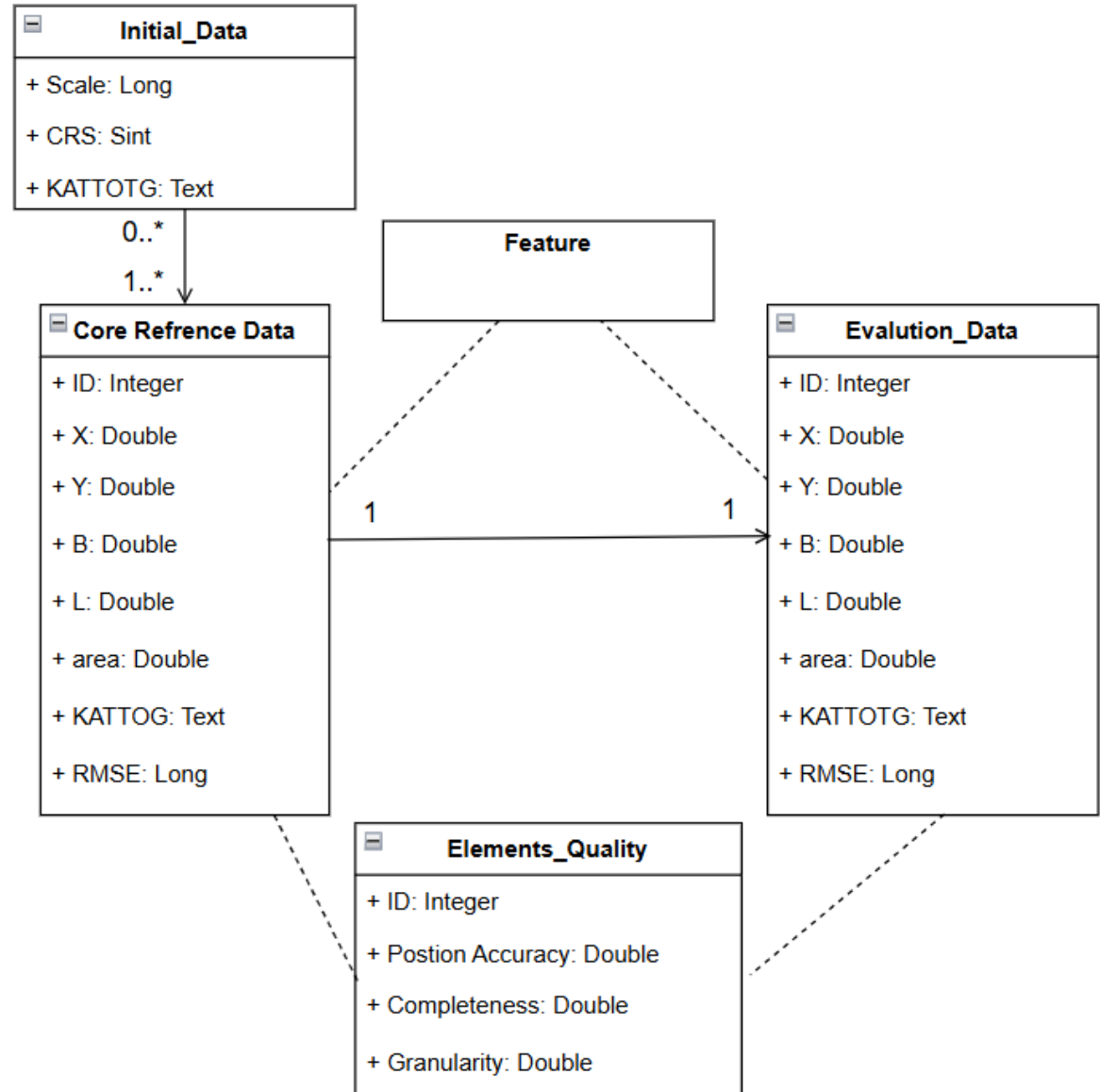
# Логічна модель бази геопросторових даних

Модель описує структуру таблиць з атрибутами та типами даних для еталонних і оцінюваних об'єктів. Встановлені зв'язки дозволяють порівнювати координати, площі та обчислювати показники якості: **позиційну точність, повноту і детальність.**

Усі поля мають визначені типи:

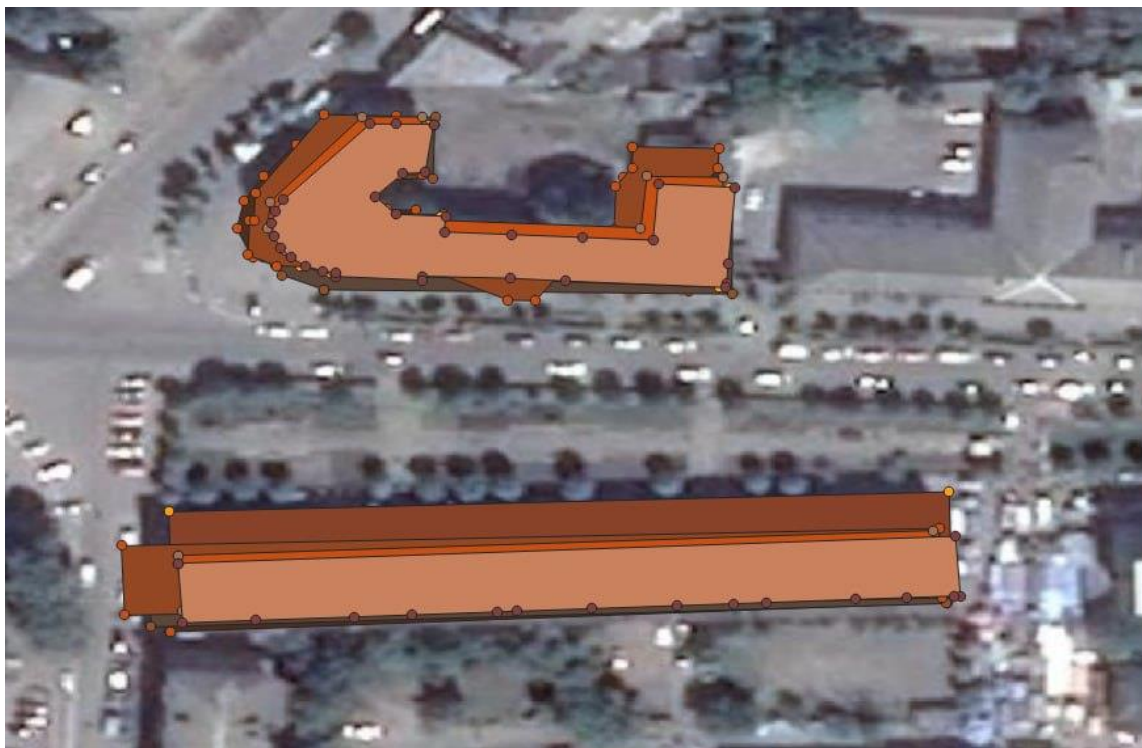
- координати — *Double*,
- площа — *Double*,
- КАТТОГТ — *Text*,
- ідентифікатори — *Integer*.

Таблиці пов'язані між собою через ID і КАТТОГ для забезпечення просторової відповідності та автоматизованого аналізу.





## Підготовка ГІС-проекту для оцінювання якості відкритих геопросторових даних



- У кожному з п'яти населених пунктів (Тульчин, Славута, Лисянка, Рогань, Пісочин) було відібрано по 30 будівель для подальшого аналізу повноти, детальності та позиційної точності.
- Відбір здійснювався на основі еталонного топографічного плану масштабу 1:2000, з урахуванням різноманітності площ і форм будівель.
- Оцифрування проводилось вручну в середовищі ArcMap, з використанням підкладок (супутникових знімків).
- Для кожного об'єкта автоматично визначались координати, площа, а також збереження здійснювалось у місцевій системі координат (МСК-05, LCS-68 тощо), яка використовувалась у межах проекту для всіх розрахунків.

# Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS

## Детальність

У межах дослідження здійснено оцінку **детальності будівельних об'єктів** з набору відкритих геопросторових даних з метою визначення відповідності вимогам масштабу **1:2000**.

Як критерій деталізації використано **граничну довжину сторони 40 м**, згідно з нормативами точності векторних моделей.

### Етапи аналізу:

- 1.Перетворення полігонів у лінійні об'єкти (Feature to Line);
- 2.Розбиття ліній на сегменти (Split Line at Vertices);
- 3.Обчислення довжин кожного сегмента;

Класифікація будівель на три групи:

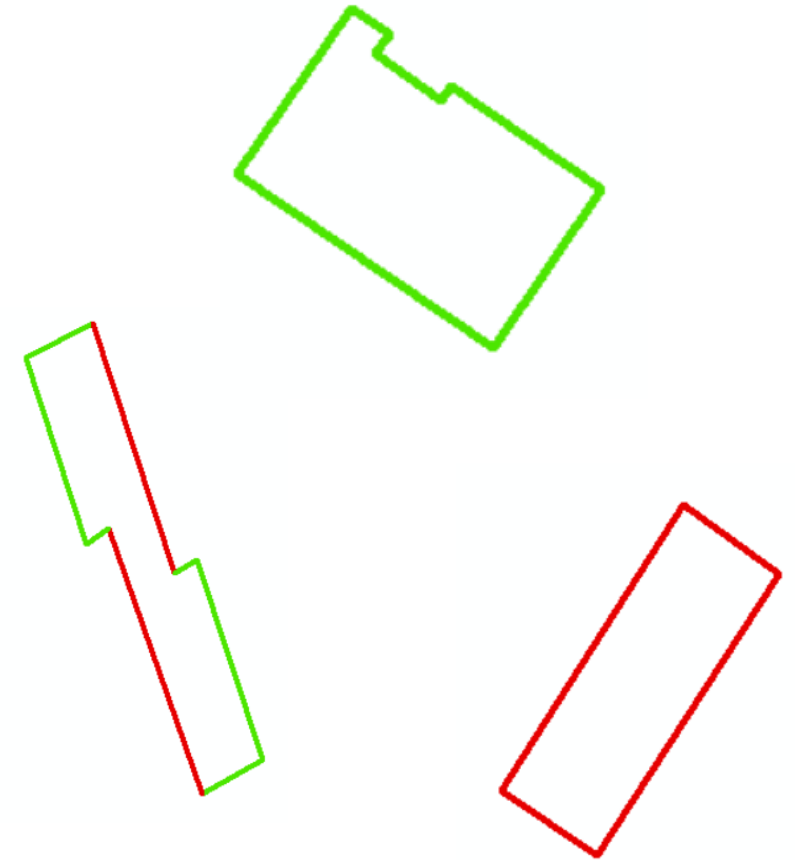
- Повністю відповідає (усі сторони  $\leq 40$  м)
- Частково відповідає (частина сторін  $> 40$  м)
- Не відповідає повністю (усі сторони  $> 40$  м)

### Візуалізація:

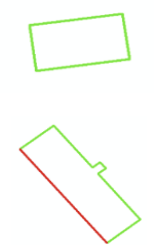
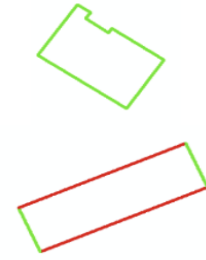
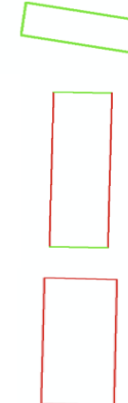
Для наочності здійснено кольорове кодування сегментів:


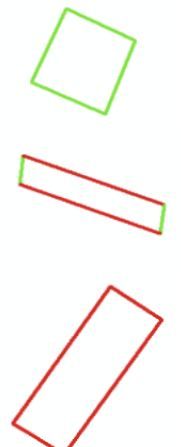
**Зелений колір** — сторони, що відповідають критерію;

**Червоний колір** — сторони, що перевищують допустиму довжину.



# Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS

Набір відкритих даних	Повністю відповідає (к-сть)	Частково відповідає (к-сть)	Не відповідає повністю (к-сть)	Візуальні приклади об'єктів
OSM	31	92	0	
Microsoft Building Footprints	83	65	0	
Google Satellite	88	59	2	

Набір відкритих даних	Повністю відповідає (к-сть)	Частково відповідає (к-сть)	Не відповідає повністю (к-сть)	Візуальні приклади об'єктів
Esri Imagery	84	63	2	
Bing Maps	86	61	2	

# Виконання оцінки якості відкритих геопросторових даних у середовищі ArcGIS

## Повнота

Набір даних	Мін. площа (м <sup>2</sup> )	Макс. площа (м <sup>2</sup> )	Орієнтовний масштаб
Microsoft Building Footprints	0,2542	28983,8289	1:500 -1:50000
OSM	7,5542	28840,9448	1:2000 -1:50000
Google Satellite	59,8087	9717,2639	1:5000 -1:50000
Esri Imagery	59,1054	9763,0554	1:5000 -1:50000
Bing Maps	45,2426	9624,4272	1:5000- 1:50000

У даному випадку повнота оцінювалась шляхом порівняння площ будівель, отриманих з різних джерел, а також аналізу розмаху (мінімальної та максимальної площ) об'єктів, що дозволяє встановити орієнтовний масштаб використання даних.

Для кожного набору даних визначались:

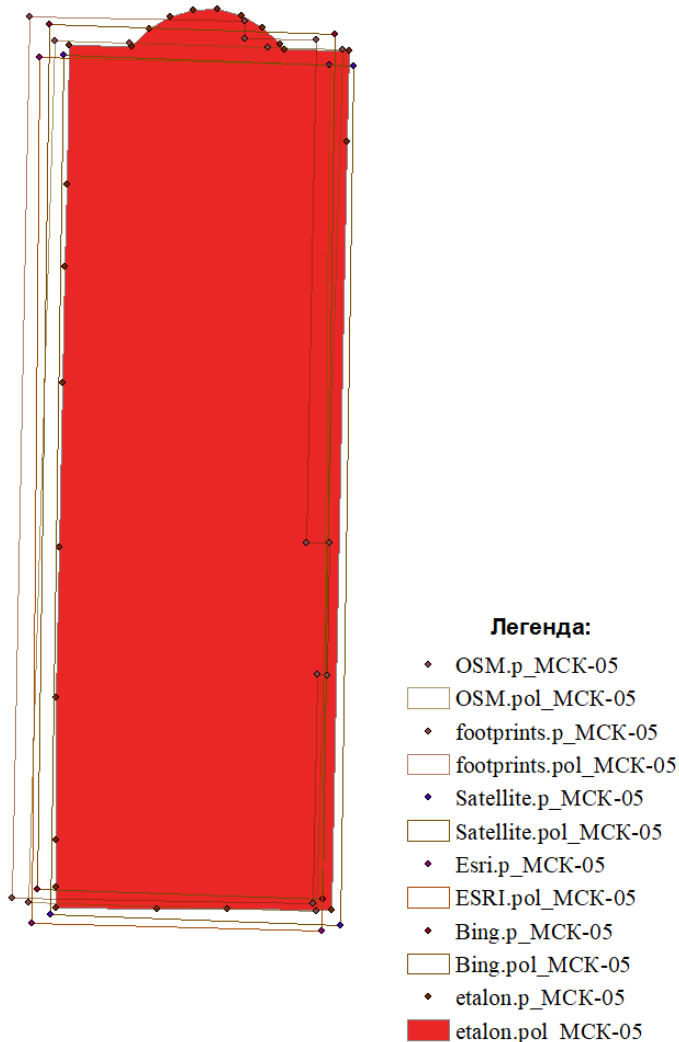
1. **Мінімальна площа об'єкта** – відображає здатність джерела фіксувати дрібні, незначні за площею будівлі.
2. **Максимальна площа об'єкта** – показує, чи були належним чином охоплені великі об'єкти.
3. **Орієнтовний масштаб**, в якому дані можуть використовуватись без суттєвих втрат детальності.

## Позиційна точність

У рамках дослідження оцінка позиційної точності не проводилась через несумісність конфігурацій об'єктів з відкритих джерел та еталонних даних.

Основні причини:

- Різна кількість вершин у контурах будівель та споруд;
- Відмінності у загальному контурі геопросторових моделей будівель та споруд;
- Ускладнення зіставлення відповідних точок моделей одночасно в усіх наборах даних, що потребує додаткового опрацювання.



# Висновок

За результатами оцінки:

**Найбільшого значення показника за повнотою** продемонстрував набір **Microsoft Building Footprints**, де мінімальна площа об'єкта становила **0,25 м<sup>2</sup>**, максимальна – **28 983,8 м<sup>2</sup>**. Це свідчить про наявність полігональних моделей об'єктів, які придатні як додатковий вихідний матеріал для топографічного плану **масштабу від 1:500**.

- **OpenStreetMap** охоплює великі за розміром об'єкти (до **28 840,9 м<sup>2</sup>**), однак мінімальна їх площа – **7,55 м<sup>2</sup>**, що обмежує повноту для щільної забудови. Придатний як додатковий вихідний матеріал для топографічного плану **масштабу від 1:2000**.
- Набори **Google, ESRI та Bing** містять об'єкти з мінімальними площами **від 45 до 59 м<sup>2</sup>**, максимальними – до **9700 м<sup>2</sup>**, що обмежує використання даних масштабом **1:5000 і дрібніше**.

За **детальністю** найкращі показники мають Google (59,1%), Bing (57,7%), Esri (56,4%) і Microsoft (56,1%) – будівлі повністю відповідають критерію (усі сторони  $\leq 40$  м).

А набір **OSM** має лише **25,2%** будівель з повною відповідністю зазначеній вимозі, решта – спрощені, хоча повністю непридатних об'єктів не зафіксовано.

Отже, **найбільш придатним для використання як додатковий вихідний матеріал під час створення цифрового топографічного плану масштабу 1:2000 є Microsoft Building Footprints**, тоді як **OSM** та супутникові джерела варто використовувати з урахуванням обмежень по детальності та мінімальних площах.

**Дякую за увагу!**

---